PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁵ :		(11) International Publication Number	wo 94/25591
C12N 15/13, C07K 15/28, A61K 39/395	A1	(43) International Publication Date:	10 November 1994 (10.11.94)

(21) International Application Number:

PCT/EP94/01442

(22) International Filing Date:

28 April 1994 (28.04.94)

(30) Priority Data:

29 April 1993 (29.04.93) EP 93201239.6 (34) Countries for which the regional or international application was filed: NL et al. 19 May 1993 (19.05.93) EP (34) Countries for which the regional or international application was filed: NL et al. 15 July 1993 (15.07.93) 93202079.5 EP (34) Countries for which the regional or international application was filed: NL et al.

- (71) Applicant (for all designated States except AU BB CA GB IE LK MN MW NZ SD US): UNILEVER N.V. [NL/NL]; Weena 455, NL-3013 AL Rotterdam (NL).
- (71) Applicant (for AU BB CA GB IE LK MN MW NZ SD only):
 UNILEVER PLC [GB/GB]; Unilever House, Blackfriars,
 London EC4 4BQ (GB).
- (71)(72) Applicants and Inventors: HAMERS, Raymond [BE/BE]; Vijversweg 15, B-1640 Sint-Genesius-Rode (BE). HAMERS-CASTERMAN, Cécile [BE/BE]; Vijversweg

15, B-1640 Sint-Genesius-Rode (BE). MUYLDERMANS, Serge, Victor, M. [BE/BE]; Brusselse Steenweg 55, B-1560 Hoeilaart (BE).

- (72) Inventors; and
- (75) Inventors/Applicants (for US only): FRENKEN, Leon, Gerardus, J. [NL/NL]; Geldersestraat 90, NL-3011 MP Rotterdam (NL). VERRIPS, Cornelis, Theodorus [NL/NL]; Hagedoorn 18, NL-3142 KB Maassluis (NL).
- (74) Common Representative: UNILEVER N.V.; Patent Division, P.O. Box 137, NL-3130 AC Vlaardingen (NL).
- (81) Designated States: AT, AU, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, ES, FI, GB, GE, HU, JP, KG, KP, KR, KZ, LK, LU, LV, MD, MG, MN, MW, NL, NO, NZ, PL, PT, RO, RU, SD, SE, SI, SK, TJ, TT, UA, US, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published

With international search report.

(54) Title: PRODUCTION OF ANTIBODIES OR (FUNCTIONALIZED) FRAGMENTS THEREOF DERIVED FROM HEAVY CHAIN IMMUNOGLOBULINS OF CAMELIDAE

(57) Abstract

A process is provided for the production of an antibody or a fragment or functionalized fragment thereof using a transformed lower eukaryotic host containing an expressible DNA sequence encoding the antibody or (functionalized) fragment thereof, wherein the antibody or (functionalized) fragment thereof is derived from a heavy chain immunoglobulin of Camelidae and is devoid of light chains, and wherein the lower eukaryotic host is a mould, preferably belonging to the genera Aspergillus or Trichoderma, or a yeast, preferably belonging to the yeast genera Saccharomyces, Kluyveromyces, Hansenula, or Pichia. The heavy chain fragment can contain at least the whole variable domain. A complementary determining region (CDR) different from the CDR belonging to the natural antibody ex Camelidae can be grafted on the framework of the variable domain of the heavy chain immunoglobulin. The catalytic antibodies can be raised in Camelidae against transition state molecules. The functionalized antibody or fragment thereof can comprise a fusion protein of both a heavy chain immunoglobulin from Camelidae or a fragment thereof and another polypeptide, e.g., an enzyme, preferably an oxido-reductase. Also provided are new products obtainable by a process as described, and compositions containing a product produced by a process as described, which composition may contain a new product as provided.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
ΑÜ	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	rr	Italy	PL	Poland
BR	Brazil	JР	Japan	PT	Portugal
BY	Belarus	KE	Кепуа	RO	Romania
CA	Canada	KG	Kyrgystan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic	SD	Sudan
CG	Congo		of Korea	SE	Sweden
CH	Switzerland	KR	Republic of Korea	SI	Slovenia
CI	Côte d'Ivoire	KZ	Kazakhstan	SK	Slovakia
CM	Cameroon	LI	Liechtenstein	SN	Senegal
CN	China	LK	Sri Lanka	TD	Chad
CS	Czechoslovakia	LU	Luxembourg	TG	Togo
CZ	Czech Republic	LV	Latvia	TJ	Tajikistan
DE	Germany	MC	Monaco	TT	Trinidad and Tobago
DK	Denmark	MD	Republic of Moldova	UA	Ukraine
ES	Spain	MG	Madagascar	US	United States of America
FI	Finland	MIL	Mali	UZ	Uzbekistan
FR	France	MN	Mongolia	VN	Vict Nam
GA	Gabon		•		

WO 94/25591

1

PCT/EP94/01442

Title: Production of antibodies or (functionalized) fragments thereof derived from heavy chain immunoglobulins of *Camelidae*

The present invention relates to a process for the production of antibodies or (functionalized) fragments thereof derived from heavy chain immunoglobulins of *Camelidae* and is partly based on research investigations carried out at the Free University of Brussels. A draft publication thereon already submitted to the periodical Nature and communicated to the present applicants by Prof. R. Hamers reads as follows.

10

15

20

25

5

FUNCTIONAL HEAVY CHAIN IMMUNOGLOBULINS IN THE CAMELIDS

Random association of V_L and V_H repertoires contributes considerably to antibody diversity (1). The diversity and the affinity are then increased by hypermutation in B-cells located in germinal centres (2). Except in the heavy chain disease (3), naturally occurring heavy chain antibodies have not been described, although antigen binding has been demonstrated for separated heavy chains (4) or cloned V_H domains (5). The presence of considerable amounts IgG like material of 100 Kd in the serum of the camel (*Camelus dromedarius*) (6) was confirmed. These molecules are composed of heavy chain dimers and are devoid of light chains. Nevertheless they bear an extensive antigen binding repertoire, a finding which questions the role of the light chains in the camel. Camel heavy chain IgGs lack the C_H1, which in one IgG class might be structurally replaced by an extended hinge. Heavy chain IgGs are a feature of all camelids. These findings open perspectives in engineering of antibodies.

By a combination of affinity chromatography on Protein A and Protein G, three quantitatively important fractions corresponding to subclasses of IgG can be isolated from the serum of camels (*Camelus dromedarius*) (Fig. 1A, lanes c-f).

One fraction (IgG₁) contains molecules of 170 Kd (Fig. 1B, lane 2) which upon reduction yield 50 Kd heavy chains and large 30 kD light chains (Fig. 1C, lane 2). The two other immunoglobulin fractions contain molecules of approximately 100 Kd

(Fig. 1B, lanes 1 and 3) which upon reduction yield only heavy chains of respectively 46 Kd (IgG₂ fraction binding only to Protein A) (Fig. 1C, lane 3) and 43 Kd (IgG₃ fraction binding to Protein A and Protein G) (Fig. 1C, lane 1). These two IgG classes appear to lack the light chain completely.

5

10

To exclude the possibility that the light chains were only weakly associated with the heavy chains and lost during the selective purification, whole serum was size fractionated by gel filtration. Coomassie blue staining of unreduced fractions revealed the sequential elution of the 170 Kd IgG₁ followed by the incompletely resolved isotypes IgG₂ and IgG₃ (90 Kd) (Fig. 1D, upper inset). Immunostaining of the same fractions after reduction confirmed that the light chains were present solely in the 50 Kd heavy chain containing fractions (Fig. 1D, lower inset).

A comparative study of old world camelids (Cameles bactrianus and Camelus dromedarius) and new world camelids (Lama pacos, Lama glama and Lama vicugna) showed that heavy chain immunoglobulins are abundant in the sera of all species examined (data not shown) and total up to 75% of the molecules binding to protein A.

The abundance of the heavy chain immunoglobulins in the serum of camelids raises

the question as to whether they bear an extensive antigen binding repertoire. This question could be answered by examining the IgG₁, IgG₂ and IgG₃ fractions from the serum of camels (*Camelus dromedarius*) with a high antitrypanosome titer (7). In radio-immunoprecipitation, purified fractions of IgG₁, IgG₂ and IgG₃ derived from infected camels were shown to bind a large number of antigens present in a ³⁵S methionine labelled trypanosome lysate (Fig. 2A), indicating an extensive repertoire complexity for the three IgG classes. Conversely, in blotting experiments, ³⁵S methionine labelled trypanosome lysate binds to SDS-PAGE separated IgG₁, IgG₂ and IgG₃ obtained from infected animals (Fig. 2B). These findings indicate that the heavy chains alone can generate an extensive repertoire and question the obligatory contribution of the light chain to the useful antibody repertoire in the camelids.

WO 94/25591 PCT/EP94/01442

3

The camelid $\gamma 2$ and $\gamma 3$ chains are considerably shorter than the normal mammalian γ or camel $\gamma 1$ chains. This would suggest that, as in the case of heavy chain disease (3), deletions have occurred in the $C_{II}1$ protein domain (8,9). To address this question, cDNA was synthesized from camel spleen mRNA and the sequences between the 5' end of the V_{II} and the $C_{II}2$ were amplified by a Polymerase Chain Reaction (PCR), and cloned. Seventeen clones presenting a different V_{II} sequence were isolated and sequenced. Their most striking feature was the complete lack of the $C_{II}1$ domain, the last framework (FR4) residues of the V_{II} region being immediately followed by the hinge (Fig. 3, lower part). The absence of the $C_{II}1$ domain clarifies two important dilemmas.

First, immunoglobulin heavy chains are normally not secreted unless the heavy chain chaperoning protein or BIP (10) has been replaced by the L chain (11), or alternatively the C_H1 domain has been deleted (3,8,9). Secondly, isolated heavy chains from mammalian immunoglobulins tend to aggregate, but are only solubilized by light chains (8,12) which bind to the C_H1 and the V_H domains (13).

14 of the 17 clones were characterized by a short hinge sequence with a length equal to that of human IgG_2 and IgG_4 (14) (Fig. 3). The other 3 had a long hinge sequence containing the 'EPK' hinge motif found in human IgG_1 and IgG_3 (14). They possess the C_H2 'APELL/P' motif also found in human IgG_1 and IgG_3 (see SEQ. ID. NO: 1-2), and which is associated with mammary transport of bovine IgG_1 (15). On basis of molecular weight, we expect the "short hinge" clones to correspond to IgG_3 and the "long hinge" clones to IgG_2 .

25

10

15

20

In the short hinge containing antibody, the extreme distance between the extremities of the V_H regions will be of the order of 80 Å corresponding to twice the size of a single domain of 40 Å $(2xV_H)$ (16). This could be a severe limitation for agglutinating, cross linking or complement fixation (17,18). In the long hinge containing immunoglobulin the absence of $C_H 1$ might be compensated by the extremely long hinge itself, composed of a 12 fold repeat of the sequence Pro-X (X=Gln, Glu, Lys) (Fig. 3 & 4). NMR (19) and molecular modelling (20) of Pro-X repeats present in

WO 94/25591 PCT/EP94/01442

the TonB protein of E. coli (X=Glu, Lys) and the membrane procyclin of trypanosomes (X=Asp, Glu) indicate that these repeated sequences function as rigid rodlike spacers with a diameter of 8 Å and a rise of 2.9 Å per residue. Assuming the same geometry, the long hinge would be 70 Å which compensates for the absence of the $C_{\rm H}1$ domain.

The binding site of heavy chain antibodies cannot form the pocket resulting from adjoining light and heavy chain V regions and the residues of the V_H which normally interact with V_L will be exposed to solvent (3,5,13). It was found that leucine at position 45 conserved in 98% of human and murine V_H sequences (14), and crucial in the V_H-V_L association (13), can be replaced by an arginine (Fig. 3, upper part). This substitution is in accordance with both the lost contact with a V_L domain and an increased solubility.

Unlike myeloma heavy chains which result mainly from C_H1 deletion in a single antibody producing cell (21) the camelid heavy chain antibodies have emerged in a normal immunological environment and it is expected that they will have undergone the selective refinement in specificity and affinity accompanying B cell maturation (1, 2). The obtention of camelid heavy chain antibodies could therefore be an invaluable asset in the development and engineering of soluble V_H domains (5) or of new immunologicals for diagnostic, therapeutic or biochemical purposes.

<u>REFERENCES</u>

- 1. Tonegawa, S. Nature 302, 575-581 (1983).
- 25 2. Jacob, J., Kelsoe, G., Rajewski, K., & Weiss, U. Nature 354, 389-392 (1991).
 - 3. Fleischman J.B., Pain R.H. & Porter R.R. Arch. Biochem. Biophys Suppl. 1, 174-180 (1962).
 - 4. Utsumi, S. & Karush, F. *Biochemistry* 3, 1329-1338 (1964).
 - 5. Ward, E.S., Güssow, D., Griffiths, A.d., Jones, P.T. & Winter G. Nature 341,
- 30 544-546 (1989).
 - 6. Ungar-Waron H., Eliase E., Gluckman A. and Trainin Z. *Isr. J. Vet. Med.* 43, 198-203 (1987).

- 7. Bajyana Songa, E., & Hamers R. Ann. Soc. Belge Méd. Trop. 68, 233-240 (1988).
- 8. Seligmann M., Mihaesco E., Preud'homme J.-L., Danon F. & Brouet J.-C. *Immun.Rev.* 48, 145-167 (1979).
- 9. Traunecker, A., Schneider, J., Kiefer, H., Karjalaien, K., *Nature* 339, 68-70 (1989).
 - 10. Henderschot L.M., Bole D., Köhler, G. & Kearney, J.F. J. Cell Biol. 104, 761-767 (1987).
 - 11. Henderschot L.M. J. Cell Biol. 111, 829-837 (1990).
- 12. Roholt O., Onoue K. & Pressman D. *Proc.Natn.Acad.Sci. USA* **51**, 173-178 10 (1964).
 - 13. Chothia, C., Novotny, J., Bruccoleri, R., Karplus, M. J. Mol. Biol. 186, 651-663 (1985).
 - 14. Kabat E.A., Wu, T.T., Reid-Miller, M., Perry H.M. & Gottesman, K.S. Sequences of Proteins of Immunological Interest 511 (U.S. Dept of Health and Human Services,
- 15 US Public Health Service, National Institutes of Health, Bethesda, 1987).
 - 15. Jackson, T., Morris, B.A, Sanders, P.G. Molec. Immun. 29, 667-676 (1992).
 - 16. Poljak R.J. et al. Proc.Natn.Acad.Sci. USA 70, 3305-3310 (1973).
 - 17. Dangl J.L., et al. EMBO J. 7, 1989-1994 (1988).
 - 18. Schneider W.P. et al. Proc. Natn. Acad. Sci USA 85, 2509-2513 (1988).
- 20 19. Evans, J.S. et al. FEBS Lett. 208, 211-216 (1986).
 - 20. Roditi, I. et al. J.Cell Biol. 108, 737-746 (1989).
 - 21. Dunnick, W., Rabbits, T.H., Milstein, C. Nucl. Acids Res., 8, 1475-1484 (1980).
 - 22. Bülow, R., Nonnengässer, C., Overath, P. Mol.Biochem.Parasitol. 32, 85-92 (1989).
- 23. Sambrook, J., Fritsch, E.F. & Maniatis, T. *Molecular Cloning: A Laboratory Manual* 2nd Edn (Cold Spring Harbor Laboratory Press, New York, 1989).
 - 24. Sastry, L et al. Proc. Natn. Acad. Sci. USA 86, 5728-5732 (1989).
 - 25. Sanger, F., Nicklen, S. & Coulson, A.R. *Proc.Natn.Acad.Sci. USA* 74, 5463-5467 (1977).
- 30 26. Klein, J. *Immunology* (Blackwell Scientific Publications, London, 1990).

Figure 1 Characterisation and purification of camel IgG classes on Protein A,

Protein G and gel filtration.

- (A) The fraction of *C. dromedarius* serum adsorbed on Protein A shows upon reduction on SDS-PAGE three heavy chain components of respectively 50, 46, and 43 Kd (bands between dots), absent in the non adsorbed fraction (lane d), and light chain components of around 30 Kd (lane c) considerably larger than rabbit light chain (lane a, rabbit IgG). The fractions adsorbed on Protein G (lane e) lack the 46 Kd heavy chain which remains in the non adsorbed fraction (lane f). Lane b contains a size marker.
- 10 (B and C) By differential adsorption and elution on Protein G and Protein A, the IgG fractions containing 43 Kd (lane 1), 46 Kd (lane 3) and 50 Kd (lanes 2) heavy chains were purified and analysed on SDS-PAGE in absence (B) or presence (C) of DTT.
- (D) Whole camel serum (0.1 ml) was fractionated by gel filtration on a Superdex 200 column using 150 mM NaCl, 50 mM sodium phosphate buffer pH 7.0 as eluent. Affinity purified IgG₂ and IgG₃ elute at the positions indicated by arrows. The fractions of interest were further analysed by SDS-PAGE with or without prior reduction. The protein contents as visualized by Coomassie blue (without reduction, upper inset) are compared with the immunoglobulins from the same fractions (after reduction with DTT, lower inset) as revealed by Western blotting with a rabbit anticamel-IgG (lower inset).

METHODS. 5 ml of C. dromedarius serum is adsorbed onto a 5 ml Protein G

Sepharose (Pharmacia) column, and washed with 20 mM phosphate buffer, pH 7.0. Upon elution with 0.15 M NaCl, 0.58 % acetic acid (pH 3.5), IgG₃ of 100 Kd is eluted which upon reduction yields heavy chains of 43 Kd (lane 1, B and C). IgG₁ of 170 Kd can subsequently be eluted with pH 2.7 buffer (0.1 M Gly-HCl). This fraction, upon reduction, yields a 50 Kd heavy chain and a broad light chain band (lane 2, C). The fraction not adsorbed on Protein G is brought on a 5 ml Protein A

30 Sepharose column. After washing and elution with 0,15 M NaCl, 0.58% acetic acid (pH 4.5) IgG₂ of 100 Kd is obtained which consists solely of 46 Kd heavy chains (lane 3, C).

Figure 2 Repertoire complexity and antigen binding capacity of camel IgG₁, IgG₂ and IgG₃ analysed by radioimmunoprecipitation (A) or Western blotting (B & C).

- (A) Serum or purified IgG fractions from healthy or *Trypanoma evansi* infected *C. dromedarius* (CATT titer 1/160 (7)) were incubated with labelled trypanosome lysate, recovered with Protein A Sepharose and analysed by SDS-PAGE. The relative counts recovered are inscribed below each lane. No trypanosome proteins bind to the Protein A or to the healthy camel immunoglobulins.
- 10 (B) 20 μg of IgG₁, IgG₂ and IgG₃ from healthy and trypanosome infected animals were separated by SDS-PAGE without prior reduction or heating. The electroblotted proteins were incubated with the labelled trypanosome lysate. The IgG₂ shows a single antigen binding component corresponding to the heavy chain immunoglobulin whereas the IgG₃ fraction appears to contain in addition two
 15 larger antigen binding components barely detectable by Ponceau Red staining (C). These are possibly Ig classes copurified as immunocomplexes present in the serum of the infected animals.

METHODS. (35S)-methionine labelled Trypanosoma evansi lysate (500,000 counts) 20 (22) was incubated (4°C, 1 hour) with 10 μl of serum or, 20 μg of IgG₁, IgG₂ or IgG₃ in 200 µl of 0.4 M NaCl, 10 mM EDTA, 10 mM Tris (pH 8.3), containing 0.1 M TLCK. 10 mg of Protein A SeDharose suspended in 200 µl of the same buffer was added (4°C, 1 hour). After washing and centrifugation, each pellet was resuspended in 75 µl SDS PAGE sample solution containing DTT, and heated for 3 min. at 25 100°C. After centrifugation, 5 µl of the supernatant was saved for radioactivity counting and the remainder analysed by SDS PAGE and fluorography. The nitrocellullose filter of the Western blot of purified fractions IgG₁, IgG₂ and IgG₃ was stained with Ponceau Red (C) or incubated with 1% ovalbumin in TST buffer (Tris 10 mM, NaCl 150 mM, Tween 0,05%) (B). The membrane was extensively washed with TST buffer and incubated for 2 hours with (35S)-labelled 30 trypanosome antigen. To avoid unspecific binding, the labelled trypanosome antigen

10

15

20

25

lysate was filtered (45 μ) and incubated with healthy camel immunoglobulin and ovalbumin adsorbed on a nitrocellulose membrane.

Figure 3 Amino acid sequences of the V_{II} framework, and hinge/ C_{II} 2 of Camelus dromedarius heavy chain immunoglobulins, compared to human (italic) V_{II} framework (subgroup III) and hinges of human IgG (14).

METHODS. Total RNA was isolated from a dromedary spleen (23). mRNA was purified with oligo T-paramagnetic beads (PolyATract-Promega). 1 μg mRNA was used for preparing double-strand cDNA (23) after an oligo-dT priming using enzymes provided by Boehringer Mannheim. 5 µg of cDNA was amplified by PCR in a 100 µl reaction mixture (10mM Tris-HCl pH 8.3, 50 mM KC1,15 mM MgCl₂, 0.01% (w/v) gelatine, 200 µM of each dNTP). 25 pmoles of each oligonucleotide of the mouse V_H (24), containing a XhoI site, and 5'-CGCCATCAAGGTACCAGT-TGA-3' (see SEQ. ID. NO: 3) were used as primers. The 3' end primer was deduced from partial sequences corresponding to y chain amino acid 296 to 288 (T.Atarhouch, C. Hamers-Casterman, G. Robinson, private communication) in which one mismatch was introduced to create a KDnI restriction site. After a round of denaturing annealing (94°C for 5 min. and 54°C for 5 min.), 2 U of Taq DNA polymerase were added, to the reaction mixture before subjecting it to 35 cycles of amplification (5). The PCR products were purified by phenol-chloroform extraction followed by HPLC (Genpak-fax column, Waters) and finally by MERMAID (BIO 101, Inc.). After these purification steps, the amplified cDNA was digested with XhoI and KpnI, and ligated into pBluescript.

The clones were sequenced by the dideoxy chain termination method (25). The sequences were translated into amino acids which allowed their assignment to well defined domains of the Ig molecule (14); see SEQ. ID. NO: 4-12

Figure 4 Schematic representation of the structural organisation of the camel immunoglobulins (adapted from 26).

On the basis of size consideration, the IgG₁ fraction possess probably the normal antibody assembly of two light and two heavy chains. IgG₃ would have a hinge comparable in size to the human IgG₁, IgG₂ and IgG₄. The two antigen binding sites

WO 94/25591 PCT/EP94/01442

9

are much closer to each other as this camel IgG lacks the $C_{11}1$ domain. In the camel IgG₂ the long hinge, being formed of Pro-X repeats (X = Glu, Gln or Lys), most likely adopt a rigid structure (19,20). This long hinge could therefore substitute the $C_{11}1$ domain and bring the two antigen binding sites of IgG₂ to normal positions.

5

--- End of Draft publication ---

Background of the invention

Already at a very early stage during evolution antibodies have been developed to 10 protect the host organisms against invading molecules or organisms. Most likely one of the earliest forms of antibodies must have been developed in Agnatha. In these primitive fishes antibodies of the IgM type consisting of heavy and lights chains have been detected. Also in many other forms of life ranging from amphibians to mammals antibodies are characterized by the feature that they consist of two heavy 15 and two light chains, although the heavy chains of the various classes of immunoglobulins are quite different. These heavy and light chains interact with each other by a number of different physical forces, but interactions between hydrophobic patches present on both the heavy and light chain are always important. The interaction between heavy and light chains exposes the complementarity determining regions (CDRs) of both chains in such a way that the immunoglobulin can bind the antigen optimally. Although individual heavy or light chains have also the capability to bind antigens (Ward et al., Nature 341 (1989) 544-546 = ref. 5 of the above given draft publication) this binding is in general much less strong than that of combined heavy and light chains.

Heavy and light chains are composed of constant and variable domains. In the organisms producing immunoglobulins in their natural state the constant domains are very important for a number of functions, but for many applications of antibodies in industrial processes and products their variable domains are sufficient. Consequently many methods have been described to produce antibody fragments.

30

One of these methods is characterized by cleavage of the antibodies with proteolytic enzymes like papain and pepsin resulting in (a) antibody fragment comprising a light

chain bound via an S-S bridge to part of a corresponding heavy chain formed by proteolytic cleavage of the heavy chain (Fab), or (b) a larger fragment of the antibody comprising two of these Fabs still connected to each other via an S-S bridge in enlargements of the heavy chain parts, indicated with F(ab)₂, respectively (see patent applications EP-A-0125023 (GENENTECH / Cabilly et al., 1984) and WO-A-93/02198 (TECH. RES. CENT. FINLAND / Teeri et al., 1993) for definitions of these abbreviations). The disadvantage of the enzymatic route is that the production of whole antibodies is expensive and the enzymatic processing increases the costs of these fragments even more. The high costs of antibody fragments block the application of these fragments in processes and products outside the pharmaceutical industry.

Another method is based on linkage on DNA level of the genes encoding (parts of) the heavy chain and the light chain. This linkage and the subsequent production of 15 these chimeric immunoglobulins in microorganisms have been described (for Fab fragments see e.g. Better et al., Science 240 (1988) 1041-1043, for F_v fragments (combination of variable fragments of the heavy chain (V_H) and light chain (V_L) still connected to each other by non-covalent binding interactions) see e.g. Skerra et al., Science 240 (1988) 1938, and for single chain F_v fragments (ScF_v; an F_v fragment in 20 which the two variable fragments are linked to each other by a linker peptide) see e.g. Bird et al., Science 242 (1988) 423-426. Provided that an appropriate signal sequence has been placed in front of the single chain V_H and V_L antibody fragment (ScF_v), these products are translocated in E. coli into the periplasmic space and can be isolated and activated using quite elaborate and costly procedures. Moreover the application of antibody fragments produced by E. coli in consumer products requires extensive purification processes to remove pyrogenic factors originating from E. coli. For this and other reasons the production of ScF, in microorganisms that are normally used in the fermentation industry, like prokaryotes as Streptomyces or Bacillus (see e.g. Wu et al. Bio/Technology 11 (1993) 71) or yeasts belonging to the 30 genera Saccharomyces (Teeri et al., 1993, supra), Kluyveromyces, Hansenula, or Pichia or moulds belonging to the genera Aspergillus or Trichoderma is preferred. However with a very few exceptions the production of ScF_v antibodies using these systems

proved to be impossible or quite poor. Although the exact reasons for the poor production are not well known, the use of linkers between the V_{II} and V_{L} chains not designed for secretion (Teeri *et al.*, 1993, *supra*) may be a reason.

Another reason may be incorrect folding of ScF_v. The frameworks and to a limited extend the CDRs of variable domains of light and heavy chains interact with each other. It has been described by Chothia et al. (J. Mol. Biol. 186 (1985) 651-663 = ref. 13 of the above given draft publication) that this interaction involves amino acids at the following positions of the variable region of the heavy chain: 35, 37, 39, 44-45, 47, 100-103 and 105 (numbering according to Kabat et al., In "Sequences of Proteins of Immunological Interest, Public Health Service, NIH, Washington DC, 1983 = ref. 14 of the above given draft publication). Especially leucine at position 45 is strongly conserved and the whole apolar side chain of this amino acid seems to be involved in the interaction with the light chain. These strong interactions may fold the ScF_v into a structure that can not be translocated in certain types of lower eukaryotes.

Thus the use of a linker in the production of ScF_v for connecting a V_H chain to a V_L chain, might negatively influence either the translocation, or the folding of such ScF_v or both.

20

30

Not prior-published European patent application 92402326.0 filed 21.08.92 (C. Casterman & R. Hamers) discloses the isolation of new animal-derived immunoglobulins devoid of light chains (also indicated as heavy chain immunoglobulins), which can especially originate from animals of the camelid family (Camelidae). This European patent specification, now publicly available as EP-A1-0 584 421, is incorporated herein by reference. These heavy chain immunoglobulins are characterized in that they comprise two heavy polypeptide chains sufficient for the formation of one or more complete antigen binding sites, whereby a complete antigen binding site means a site which will alone allow the recognition and complete binding of an antigen, which can be verified by any known method regarding the testing of the binding affinity. The European patent specification further discloses methods for

WO 94/25591 PCT/EP94/01442

isolating these heavy chain immunoglobulins from the serum of *Camelidae* and details of the chemical structure of these heavy chain immunoglobulins. It also indicates that these heavy chain immunoglobulins and derivatives thereof can be made by using recombinant DNA technology in both prokaryotes and eukaryotes. The present invention relates to a further development of the work disclosed in that prior-filed but not prior-published European specification.

Due to the absence of light chains in most of the immunoglobulins of *Camelidae* such linkers are not necessary, thereby avoiding the above-mentioned potential problems.

10

As described above in the draft publication for Nature, now publicly available as Nature 363 (3 June 1993) 446-448, and in the not prior-published European patent application 92402326.0 (supra) it was surprisingly found that the majority of the protein A-binding immunoglobulins of Camelidae consists just of two heavy chains and that these heavy chains are quite different from common forms of heavy chains, as the C_H1 domain is replaced by a long or short hinge (indicated for IgG₂ and IgG₃, respectively, in Figure 4 of the above given draft publication for Nature). Moreover these heavy chains have a number of other features that make them remarkably different from the heavy chains of common immunoglobulins.

One of the most significant features is that they contain quite different amino acid residues at those positions involved in binding to the light chain, which amino acids are highly conserved in common immunoglobulins consisting of two heavy and two light chains (see Table 1 and SEQ. ID. NO: 13-31).

Table 1 Comparison af amino acid sequences of various immunoglobulins Alignment of a number of V₁₁ regions of Camel heavy chain antibodies compared with those of mouse (M, top line) and human (H, second line). Framework fragments are indicated in capitals, CDR fragments in small print; see SEQ. ID. NO: 13-31 for sequences indicated by M, H, 1, 2, 3, 7, 9, 11, 13, 16, 17, 18, 19, 20, 21, 24, 25, 27, 29, respectively.

```
50
10
           EVKLVESGGG LVQPGGSLRL SCATSGFTFS dfyme..WVR QPPGKRLEWI
       h
          EVQLVESGGG LVQPGGSLRL SCAASGFTFS syams..WVR QAPGKGLEWV
    caml
           ......GG SVQAGGSLRL SCAASGYSNC pltws..WYR QFPGTEREFV
    cam2
          DVQLVASGGG SVQAGGSLRL SCTASGDSFS rfams..WFR QAPGKECELV
    cam3
           .....GG SVOTGGSLRL SCAVSGFSFS tscma..WFR QASGKQREGV
15
    cam7
           ......GG SVQGGGSLRL SCAISGYTYG sfcmq..WFR EGPGKEREGI
    cam9
           .....GG SVQAGGSLTL SCVYTNDTGT ...mq..WFR QAPGKECERV
   cam11
           .....GG SVQAGGSLRL SCNVSGSPSS tyclg..WFR QAPGREREGV
   cam13
           ......GG SVEAGGSLRL SCTASGYVSS ...ma..WFR QVPGQEREGV
   cam16
           ........GG SAQAGGSLRL SCAAHGIPLN gyyia..WFR QAPGKGREGV
20
   cam17
           ......GG SVQPGGSLTL SCTVSGATYS dysig..WIR QAPGKDREVV
   cam18
           ......GG SVQAGGSLRL SCTGSGFPYS tfclg..WFR QAPGKEREGV
   cam19
           ......GG SVQAGGSLRL SCAASDYTIT dycma..WFR QAPGKERELV
   cam20
           .....GG SVQVGGSLRL SCVASTHTDS stcig..WFR QAPGKEREGV
   cam21
           ......GG SVQVGGSLKL SCKISGGTPD rvpkslaWFR QAPEKEREGI
   cam24
           ......GG SVQAGGSLRL SCNVSGSPSS tyclg..WFR QAPGKEREGV
   cam25
           ......GG SVQTGGSLRL SCEISGLTFD dsdvg..WYR QAPGDECKLV
   cam27
           .....GG SVQAGGSLRL SCASSSKYMP ctydmt.WYR QAPGKEREFV
   cam29
           .....exxGG SVQAGGSLRL SCVASGFNFE tsrma..WYR QTPGNVCELV
30
           51
                                                               100
          A..asrnkan dytteysasv kgRFIVSRDT SQSILYLQMN ALRAEDTAIY
       \mathfrak{m}
       h
          S..xisxktd ggxtyyadsv kgRFTISRDN SKNTLYLQMN SLRAEDTAVY
    cam1
          S..smd...p dgntkytysv kgRFTMSRGS TEYTVFLQMD NLKPEDTAMY
35
    cam2
          S..siq...s ngrtteadsv qgRFTISRDN SRNTVYLQMN SLKPEDTAVY
          Aainsgggrt yyntyvaesv kgRFAISQDN AKTTVYLDMN NLTPEDTATY
    cam3
          A..tiln..g gtntyyadsv kgRFTISQDS TLKTMYLLMN NLKPEDTGTY
    cam7
    cam9
          A..hit...p dgmtfidepv kgRFTISRDN AQKTLSLRMN SLRPEDTAVY
   cam11
          T..aint..d gsiiyaadsv kgRFTISQDT AKETVHLQMN NLQPEDTATY
          A..fvqt..a dnsalygdsv kgRFTISHDN AKNTLYLQMR NLQPDDTGVY
   cam13
   cam16
          A..ting..g rdvtyyadsv tgRFTISRDS PKNTVYLQMN SLKPEDTAIY
   cam17
          A..aant..g atskfyvdfv kgRFTISQDN AKNTVYLQMS FLKPEDTAIY
   cam18
          A..gins..a ggntyyadav kgRFTISQGN AKNTVFLQMD NLKPEDTAIY
          A.aiqvvrsd trltdyadsv kgRFTISQGN TKNTVNLQMN SLTPEDTAIY
   cam19
45
   cam20
          A..siyf..g dggtnyrdsv kgRFTISQLN AQNTVYLQMN SLKPEDSAMY
          A..vlst..k dgktfyadsv kgRFTIFLDN DKTTFSLQLD RLNPEDTADY
   cam21
          T..aint..d gsviyaadsv kgRFTISQDT AKKTVYLQMN NLQPEDTATY
   cam24
   cam25
          Sgilsdgtpy tksgdyaesv rgRVTISRDN AKNMIYLQMN DLKPEDTAMY
          S..sin...i dgkttyadsv kgRFTISQDS AKNTVYLQMN SLKPEDTAMY
   cam27
50
   cam29
          S..siy...s dgktyyvdrm kgRFTISREN AKNTLYLQLS GLKPEDTAMY
```

Table 1 (Cont.) Comparison af amino acid sequences of various immunoglobulins Alignment of a number of V_H regions of Camel heavy chain antibodies compared with those of mouse (M, top line) and human (H, second line). Framework fragments are indicated in capitals, CDR fragments in small print; see SEQ. ID. NO: 13-31 for sequences indicated by M, H, 1, 2, 3, 7, 9, 11, 13, 16, 17, 18, 19, 20, 21, 24, 25, 27, 29, respectively.

```
101
                                                  139
10
          YCARdyygss .....y.. f.....dvWG AGTTVTVSS
          YCARxxxxx xxxxxyyyyh x....fdyWG QGTLVTVSS
       h
    caml
          YCKTalqpgg ycgygx.....clWG QGTQVTVSS
    cam2
          YCGAvslmdr isqh......gcRG QGTQVTVSL
    cam3
          YCAAvpahlg pgaildlkky .....kyWG QGTQVTVSS
15
    cam7
          YCAAelsggs celpllf.......dyWG QGTQVTVSS
    cam9
          YCAAdwkywt cgaqtggyf. .....gqWG QGAQVTVSS
   cam11
          YCAArltemg acdarwatla trtfaynyWG QGTQVTVSS
          YCAAqkkdrt rwaeprew.....nnWG QGTQVTASS
   cam13
   cam16
          FCAAgsrfss pvgstsrles .sdy..nyWG QGIQVTASS
20
          YCAAadpsiy ysilxiey......kyWG QGTQVTVSS
   cam17
          YCAAdspcym ptmpappird sfgw..ddFG QGTQVTVSS
   cam18
   cam19
          SCAAtssfyw ycttapy.....nvWG QGTQVTVSS
   cam20
          YCAIteiewy gcnlrttf......trWG QGTQVTVSS
   cam21
          YCAAnqlagg wyldpnywls vgay..aiWG QGTHVTVSS
25
   cam24
          YCAArltemg acdarwatla trtfaynyWG RGTQVTVSS
   cam25
          YCAVdgwtrk eggiglpwsv qcedgynyWG QGTQVTVSS
          YCKIdsypch ll......dvWG QGTQVTVSS
   cam27
   cam29
          YCAPveypia dmcs.....ryGD PGTQVTVSS
30
```

For example, according to Pessi et al. (1993) a subdomain portion of a V_H region of common antibodies (containing both heavy chains and light chains) is sufficient to direct its folding, provided that a cognate V_L moiety is present. Thus it might be expected from literature on the common antibodies that without V_L chains proper folding of heavy chains cannot be achieved. A striking difference between the common antibodies and the *Camelidae*-derived heavy chain antibodies is, that the highly conserved apolar amino acid <u>leucine</u> (L) at place 45 present in common antibodies is replaced in most of the *Camelidae*-derived heavy chain antibodies by the charged amino acid <u>arginine</u> (R), thereby preventing binding of the variable region of the heavy chain to that of the light chains.

Another remarkable feature is that one of the CDRs of the heavy chains of this type of immunoglobulins from *Camelidae*, CDR3, is often much longer than the

corresponding CDR3 of common heavy chains. Besides the two conserved cysteines forming a disulphide bridge in common V_H fragments, the *Camelidae* V_H fragments often contain two additional cysteine residues, one of which often is present in CDR3.

According to the present inventors these features indicate that CDR3 may play an important role in the binding of antigens by these heavy chain antibodies and can compensate for the absence of light chains (also containing CDRs) in binding of antigens by immunoglobulins in *Camelidae*.

Thus, as the heavy chains of *Camelidae* do not have special features for interacting with corresponding light chains (which are absent), these heavy chains are very different from common heavy chains of immunoglobulins and seem intrinsically more suitable for secretion by prokaryotic and lower eukaryotic cells.

The present inventors realized that these features make both intact heavy chain immunoglobulins of *Camelidae* and fragments thereof very attractive for their production by microorganisms. The same holds for derivatives thereof including functionalized fragments. In this specification the term "functionalized fragment" is used for indicating an antibody or fragment thereof to which one or more functional groups, including enzymes and other binding polypeptides, are attached resulting in fusion products of such antibody fragment with another biofunctional molecule.

Summary of the invention

In a broad sense the invention provides a process for the production of an antibody or a fragment or functionalized fragment thereof using a transformed lower eukaryotic host containing an expressible DNA sequence encoding the antibody or (functionalized) fragment thereof, wherein the antibody or (functionalized) fragment thereof is derived from a heavy chain immunoglobulin of *Camelidae* and is devoid of light chains, and wherein the lower eukaryotic host is a mould or a yeast. Thus the lower eukaryotic host can be a mould, e.g. belonging to the genera *Aspergillus* or *Trichoderma*, or a yeast, preferably belonging to the yeast genera *Saccharomyces*, *Kluyveromcyes*, *Hansenula*, or *Pichia*. Preferably the fragments still contain the whole variable domain of these heavy chains.

15

20

25

PCT/EP94/01442

The invention also provides methods to produce such heavy chain immunoglobulins or (functionalized) fragments thereof in which methods the framework or the CDRs of these heavy chains are modified by random or directed mutagenesis in such a way that the mutated heavy chain is optimized for secretion by the host microorganism into the fermentation medium.

Another embodiment of the invention is that CDRs can be grafted on these optimized frameworks (compare grafting of CDRs on human immunoglobulins as described by e.g. Jones et al., Nature 321 (1986) 522). These CDRs can be obtained from common antibodies or they may originate from heavy chain immunoglobulins

of Camelidae. The binding properties may be optimized by random or directed mutagenesis. Thus in a process according to the invention an antibody or (functionalized) fragment thereof derived from a heavy chain immunoglobulin of Camelidae can be produced which comprises a CDR different from the CDR belonging to the natural antibody ex Camelidae which is grafted on the framework of the variable domain of the heavy chain immunoglobulin ex Camelidae.

The invention also provides a method for the microbiological production of catalytic antibodies. These antibodies are preferably raised in *Camelidae* against transition state molecules following procedures similar to the one described by Lerner *et al.*, Science 252 (1991) 659-667. Using random or site-directed mutagenesis such catalytic antibodies or fragments thereof can be modified in such a way that the

catalytic antibodies or fragments thereof can be modified in such a way that the catalytic activity of these (functionalized) antibodies or fragments can be further improved.

For preparing modified heavy chain antibodies a process according to the invention is provided, in which the DNA sequence encodes a modified heavy chain immunoglobulin or a (functionalized) fragment thereof derived from *Camelidae* and being devoid of light chains, and is made by random or directed mutagenesis or both.

Thus the resulting immunoglobulin or (functionalized) fragment thereof is modified such that

- it is better adapted for production by the host cell, or
- it is optimized for secretion by the lower eukaryotic host into the fermentation medium, or
 - its binding properties (k_{on} and k_{off}) are optimized, or

- its catalytic activity is improved, or
- it has acquired a metal chelating activity, or
- its physical stability is improved.
- Another particular embodiment of the present invention relates to genes encoding fusion proteins consisting of both a heavy chain immunoglobulin from Camelidae or part thereof and a second protein or another polypeptide, e.g. an enzyme, in particular an oxido-reductase, and to expression products of such genes. By means of the heavy chain immunoglobulin (fragment) the protein or enzyme can be guided to a target thereby increasing the local efficiency of the protein or enzyme significantly. Thus according to this embodiment of the invention a process is provided, in which the functionalized antibody or fragment thereof comprises a fusion protein of both a heavy chain immunoglobulin from Camelidae or a fragment thereof and another polypeptide, e.g. an enzyme, preferably an oxido-reductase.

As a result of a process according to the invention known products may be produced, e.g. antibodies also produced by *Camelidae*, but many of the possible products will be new products, thus the invention also provides new products obtainable by a process according to the invention.

The products so produced can be used in compositions for various applications.

Therefore, the invention also relates to compositions containing a product produced by a process according to the invention. This holds for both old products and new products.

25 Brief Description of the Figures

Figures 1-4 were already described above in the draft publication.

Figure 1 Characterisation and purification of camel IgG classes on Protein A, Protein G and gel filtration.

Figure 2 Repertoire complexity and antigen binding capacity of camel IgG₁,

IgG₂ and IgG₃ analysed by radioimmunoprecipitation (A) or

Western blotting (B & C).

	Figure 3	Amino acid sequences of the V _{II} framework, and hinge/C _{II} 2 of
		Camelus dromedarius heavy chain immunoglobulins, compared to
		human (italic) V _{II} framework (subgroup III) and hinges of human
		IgG (14); see SEQ. ID. NO: 4-12.
5	Figure 4	Schematic representation of the structural organisation of the camel
		immunoglobulins (adapted from 26).
	Figure 5	DNA and amino acid sequences of the Camel V _{II} fragments fol-
		lowed by the Flag sequence as present in pB03 (Figure 5A), pB09
		(Figure 5B) and pB24 (Figure 5C); see SEQ. ID. NO: 32-37.
10	Figure 6	Nucleotide sequence of synthetic DNA fragment cloned into
		pEMBL9 (Example 1); see SEQ. ID. NO: 38-41.
	Figure 7	Schematic drawing of plasmid pUR4423
	Figure 8	Schematic drawing of plasmid pUR4426
	Figure 9	Schematic drawing of plasmid pUR2778
15	Figure 10	Schematic drawing of plasmid pUR4429
	Figure 11	Schematic drawing of plasmid pUR4430
	Figure 12	Schematic drawing of plasmid pUR4445
	Figure 13	Schematic drawing of plasmid pUR4446
	Figure 14	Schematic drawing of plasmid pUR4447
20	Figure 15	Schematic drawing of plasmid pUR4451
	Figure 16	Schematic drawing of plasmid pUR4453
	Figure 17	Schematic drawings of plasmids pUR4437 and pUR4438
	Figure 18	Schematic drawings of plasmids pUR4439 and pUR4440
	Figure 19	Nucleotide sequence of synthetic DNA fragment cloned into
25		pEMBL9 (Example 6); see SEQ. ID. NO: 42-45.
	Figure 20	Schematic drawing of plasmid pAW14B.
	Figure 21	Western blot analysis of culture medium of S. cerevisiae trans;
		formants containing pUR4423M (see A) or pUR4425M (see B).
		Samples were taken after 24 (see 1) or 48 hours (see 2). For
30		pUR4425M two bands were found due to glycosylation of the
		antibody fragment.

Detailed description of the invention

The present invention relates to the production of antibodies or (functionalized) fragments thereof derived from heavy chain immunoglobulins of *Camelidae* by eukaryotes, more in particular by lower eukaryotes such as yeasts and fungi.

Therefore, mRNA encoding immunoglobulins of *Camelidae* was isolated and transcribed into cDNA according to the procedures described in the above given draft publication and not prior-published European patent application 92402326.0. In each case primers for the PCR reaction directed to the N-terminus of the V_H domain and PCR primers that either hybridize with the C-terminal regions of the V_H domain or with the short or large hinge regions as described in the above given draft publication, or with the C-terminal region of the C_H2 or C_H3 domains can be used. In this way structural genes can be obtained encoding the following fragments of heavy chain immunoglobulins of *Camelidae* (Table 2).

15

Table 2. The various forms of immunoglobulins of *Camelidae* that can be expressed in microorganisms.

- a. the variable domain of a heavy chain;
- 20 b. the variable domain and the short hinge of a heavy chain;
 - c. the variable domain and the long hinge of a heavy chain;
 - d. the variable domain, the $C_H 2$ domain, and either the short or long hinge of a heavy chain;
 - e. a complete heavy chain, including either the short or long hinge.

25

30

According to procedures described in detail in the Examples these cDNAs can be integrated into expression vectors.

Known expression vectors for Saccharomyces, Kluyveromcyes, Hansenula, Pichia and Aspergillus can be used for incorporating a cDNA or a recombinant DNA according to the invention. The resulting vectors contain the following sequences that are required for expression: (a) a constitutive, or preferably an inducible, promoter; (b) a leader or signal sequence; (c) one of the structural genes as described in Table 2

and (d) a terminator. If the vector is an episomal vector, it preferably comprises an origin of replication as well as a selection marker, preferably a food grade selection marker, (EP-A-487159, UNILEVER / Leenhouts et al.). If the vector is an integration vector, then it preferably comprises sequences that ensure integration and a selection marker in addition to the sequences required for expression of the structural gene encoding a form of the heavy chain immunoglobulin of *Camelidae* or derivatives thereof. The preferred sequences for integration are sequences encoding ribosomal DNA (WO 91/00920, 1991, UNILEVER / Giuseppin et al.) whereas the selection marker will be preferably a food grade marker.

10 For Saccharomyces the preferred inducible promoter is the GAL7 promoter (EP-A-0255153, UNILEVER / Fellinger et al.); for Kluyveromyces the preferred inducible promoter is the inulinase promoter (not yet published EP application 92203932.6, UNILEVER / Toschka & Verbakel, which is incorporated herein by reference); for Hansenula or Pichia the preferred inducible promoter is the methanol-oxidase

promoter (Sierkstra et al., Current Genetics 19 (1991) 81-87) and for Aspergillus the preferred inducible promoter is the endo-xylanase promoter (not prior-published PCT application PCT/EP 92/02896, UNILEVER / Gouka et al., now publicly available as WO-A-93/12237, which is incorporated herein by reference).

To achieve efficient <u>secretion</u> of the heavy chain immunoglobulin or parts thereof the leader (secretion) sequences of the following proteins are preferred: invertase and α-factor for Saccharomyces, inulinase for Kluyveromyces, invertase for Hansenula or Pichia (Sierkstra et al., 1991 supra) and either glucoamylase or xylanase for Aspergillus (not prior-published PCT application WO-A-93/12237, supra). As foodgrade selection markers, genes encoding anabolic functions like the leucine2 and tryptophan3 are preferred (Giuseppin et al. 1991 supra). The present invertion

tryptophan3 are preferred (Giuseppin et al. 1991, supra). The present invention describes the heterologous production of (functionalized) derivatives or fragments of immunoglobulins in a microorganism, which immunoglobulins in nature occur not as a composite of heavy chains and light chains, but only as a composite of heavy chains. Although the secretion mechanism of mammals and microorganisms is quite similar, in details there are differences that are important for developing industrial

processes.

30

To obtain frameworks of the heavy chain immunoglobulins, that are optimally secreted by lower eukaryotes, genes encoding several different heavy chains can be cloned into the coat protein of bacteriophages and subsequently the frameworks of these heavy chain immunoglobulins can be mutated using known PCR technology, e.g. Zhou et al., (1991). Subsequently the mutated genes can be been cloned in Saccharomyces and Aspergillus and the secretion of the mutated genes can be compared with the wild type genes. In this way frameworks optimized for secretion may be selected.

Alternatively these structural genes can be linked to the cell wall anchoring part of cell wall proteins, preferably GPI-linked cell wall proteins of lower eukaryotes, which result in the expression of a chimeric protein on the cell wall of these lower eukaryotes (not prior-published EP application 92202080.5, UNILEVER / Klis et al., now publicly available as International (PCT) patent application WO-A-94/01567, which is incorporated herein by reference).

Both methods have the advantage that the binding parts of the immunoglobulins are well exposed to the surrounding of the cell, microorganism, or phage and therefore can bind antigens optimally. By changing the external conditions the binding rates and dissociation rates of this binding reaction can be influenced. Therefore, these systems are very suitable to select for mutated immunoglobulins that have different binding properties. The mutation of the immunoglobulins can either be obtained by random mutagenesis, or directed mutagenesis based on extensive molecular modelling and molecular dynamical studies.

mRNAs encoding heavy chains of immunoglobulins raised in *Camelidae* against transition state molecules (Lerner et al., 1991 supra) can be obtained using standard techniques. The structural genes encoding various forms of immunoglobulins according to the invention as summarized in Table 2 can be cloned into the coat protein of bacteriophages or as fusion with the anchoring part of cell wall proteins and can be tested on the catalytic property. In this way immunoglobulins or parts thereof having catalytic properties can be determined and selected. Genes encoding these selected immunoglobulins or parts thereof can be mutated as described before and recloned in bacteriophages, but preferably cloned as chimeric cell wall bound catalysts in lower eukaryotes. By performing appropriate catalytic assays, catalytic

WO 94/25591

22

immunoglobulins or parts thereof with improved catalytic properties can be determined and selected using standard techniques.

An important application of antibodies, especially outside the pharmaceutical industry, will be chimeric proteins consisting of the binding part of antibodies and enzymes. In this way catalytic biomolecules can be designed that have two binding properties, one of the enzyme and the other of the antibody. This can result in enzymes that have superior activity. This can be illustrated with the following examples:

- a. If the substrate of the enzymic reaction is produced by an organism or an enzyme is recognized by the binding domain of the antibody, the local concentration of the substrate will be much higher than for enzymes lacking this binding domain and consequently the enzymic reaction will be improved. In fact this is a mimic of vectorial metabolism in cells (compare e.g. Mitchell, (1979) Science 206 1148-1159);
- 15 b. If the substrate of the enzymic reaction is converted into a molecule that kills organisms, then the efficiency and specificity of killing can be increased significantly if the enzyme is equipped with an antibody binding domain that recognizes the target organism (e.g. compare Takahashi et al., (1993) Science 259 1460-1463);

20

10

The invention will be illustrated by the following Examples without being limited thereto. In previously filed Unilever patent specifications several expression vectors were described, e.g. for the yeasts S. cerevisiae, Kluyveromyces, and Hansenula, and the mould Aspergillus. Examples of these publications are EP-A-0173378

25 (UNILEVER / Ledeboer et al.), EP-A-0255153, supra, and PCT applications WO-A-91/19782 (UNILEVER / van Gorcom et al.) and (not prior-published) WO-A-93/12237, supra. The genes encoding antibodies or (functionalized) fragments thereof according to the invention can be incorporated into the earlier described expression vectors or derivatives thereof using procedures well known to a skilled person in the art. All techniques used for the manipulation and analysis of nucleic 30

acid materials were performed essentially as described in Sambrook et al. (1989)

(see also ref. 23 of the above given draft publication), except where indicated otherwise.

In the description of the Examples the following endonuclease restriction sites are used:

5	AflII	CITTAAG	Mlu1	AICGCGT
	BspHI	TICATGA	Ncol	CICATGG
	BspHI	T↓CATGA	Not	GCIGGCCGC
	BstEII	GIGTNACC	NruI	TCG↓CGA
	Eagl	CIGGCCG	Sall	GITCGAC
10	<i>Eco</i> RI	GIAATTC	Xhol	CITCGAG
	<i>Hin</i> dIII	AJAGCTT	BbsI	GAAGAC(N) ₂ ! CTTCTG(N') ₆ !

Example 1 Construction of cassettes encoding V_{II} fragments originating from Camelidae.

For the production of V_{II} fragments originating from Camelidae, the antibody gene fragments were isolated and cloned as described above in the draft publication. The thus obtained gene fragments encode the V_H region, a short or a long hinge region and about 14 amino acids of the C_H2 region. By using standard molecular biological techniques (e.g. PCR technology), the V_H gene fragments could be subcloned and equipped at their 5'-ends with a gene fragment encoding the pelB signal sequence and at their 3'-ends with a gene fragment encoding the Flag tail (13 amino acids). Three of these clones were named pB3, pB9 and pB24 and were deposited at the Centraal Bureau voor Schimmelcultures, Baarn on 20 April 1993 with deposition numbers: CBS 270.93, CBS 271.93 and CBS 272.93, respectively. The DNA and amino acid sequences of the Camelidae-V_{II} fragments followed by the Flag sequence are presented in Figure 5(A-C); see SEQ. ID. NO: 32-37.

1.1 Construction of pUR4421

15

20

25

For the construction of yeast expression plasmids encoding the V_H fragments 30 preceded by the invertase (=SUC2) signal sequence, the α -mating factor preproWO 94/25591 PCT/EP94/01442

sequence, or the inulinase signal sequence and followed by either nothing, or a Myc tail or Flag tail, the constructs described below can be prepared.

The multiple cloning site of plasmid pEMBL9 (Denthe et al., 1983) (ranging from the EcoRI to the HindIII site) was replaced by a synthetic DNA fragment having the nucleotide sequence as indicated in Figure 6; see SEQ. ID. NO: 38-41. The 5'-part of this nucleotide sequence comprises an EagI site, the first 4 codons of the Camelidae V_{II} gene fragment and a XhoI site coinciding with codons 5 and 6. The 3'-part comprises the last 5 codons of the Camelidae V_{II} gene (encoding VTVSS; see SEQ. ID. NO: 47) part of which coincides partially with a BstEII site), eleven codons of the Myc tail, and an EcoRI site. The EcoRI site, originally present in pEMBL9, is not functional any more, because the 5'- end of the nucleotide sequence contains AATTT instead of AATTC, indicated in Figure 6 as "(EcoRI)". The resulting plasmid is called pUR4421.

15 1.2 Constructs with Flag tail.

After digesting the plasmid pB3 with XhoI and EcoRI, a DNA fragment of approximately 425 bp was isolated from agarose gel. This fragment codes for a truncated V_H-Flag fragment, missing the first 5 amino acids of the Camelidae V_H. The obtained fragment can be cloned into pUR4421. To this end plasmid pUR4421 can be digested with XhoI and EcoRI, after which the about 4 kb vector fragment can be isolated from an agarose gel. Ligation with the about 425 bp fragment will result in plasmid pUR4421-03F.

1.3 Constructs with Myc tail.

- After digesting the plasmid pB3 with XhoI and BstEII, a DNA fragment of approximately 365 bp was isolated from agarose gel. This fragment codes for a truncated V_{II} fragment, missing both the first 4 (QVKL; see SEQ. ID. NO: 46) and the last 5 (VTVSS; see SEQ. ID. NO: 47) amino acids of the Camelidae V_H fragment.
- 30 The obtained fragment was cloned into pUR4421. To this end plasmid pUR4421 was digested with *XhoI* and *BstEII*, after which the about 4 kb vector fragment was isolated from an agarose gel. Ligation with the about 365 bp fragment resulted in

plasmid pUR4421-03M, in which the gene encoding the V_{II} fragment is reconstituted.

1.4 Constructs encoding V_{II} only.

5 Upon digesting pUR4421-03M or pUR4421-03F with *BstEII* and *HindIII*, the vector fragments of about 4.4 kb can be isolated from agarose gel and religated in the presence of a synthetic linker peptide having the following sequence:

BSTEII HINDIII

GTCACCGTCTCCTCATAATGA

10 GCAGAGGAGTATTACTTCGA (see SEQ. ID. NO: 48-49).

In the thus obtained plasmid, pUR4421-03, the Myc tail or Flag tail sequences are removed and the $V_{\rm H}$ gene fragment is directly followed by a stop codon.

1.5 Other constructs.

other by a peptide linker.

- After isolating the gene fragments encoding V_{II}-hinge-C_{II}2 fragments as described above in the draft publication, or encoding the intact heavy chain immunoglobulin, it is possible, e.g. by using PCR technology, to introduce an appropriate restriction enzyme recognition site (e.g. EcoRI or HindIII) downstream of the hinge region, downstream of the C_{II}2 region, or downstream of the total gene. Upon isolating a
- 20 XhoI-EcoRI or XhoI-HindIII fragment encoding the V_H fragment with a C-terminal extension, the fragment can be cloned into pUR4421 digested with the same restriction enzymes.

In analogy with the construction of pUR4421-03, a number of other constructs can be produced encoding functionalized heavy chain fragments in which a second polypeptide is fused to the C-terminal part of the V_H fragment. Optionally, the V_H fragment and the second polypeptide, e.g. an enzyme, might be connected to each

To this end either the *Bst*EII-*Hin*dIII fragment or the *Bst*EII-*Eco*RI fragment of either pUR4421-03F or pUR4421-03M has to be replaced by another *Bst*EII-*Hin*dIII or *Bst*EII-*Eco*RI fragment. The latter new fragment should code for the last amino acids (VTVSS, see SEO.ID. NO: 47) of the V_{II} fragment, optionally for a linker peptide, and for the polypeptide of interest e.g. an enzyme. Obviously, the introduction of the DNA fragment should result in an in frame fusion between the

 V_{II} gene fragment and the other DNA sequence encoding the polypeptide of interest.

Alternatively, it is possible to replace the EagI-XhoI fragment of pUR4421-03 with another DNA fragment, coding for a polypeptide of interest, optionally for a peptide linker, and for the first 4 (QVKL, see SEQ.ID. NO: 46) amino acids of the V_H fragment, resulting in an in frame fusion with the remaining part of the V_H fragment. In this way, it is possible to construct genes encoding functionalized V_H fragments in which the second polypeptide is fused at the N-terminal part of the V_H fragment, optionally via a peptide linker.

Obviously, it is also possible to construct genes encoding functionalized V_H fragments having a polypeptide fused to the N-terminal as well as fused to the C-terminal end, by combining the above described construction routes.

The polypeptides used to functionalize the V_{II} fragments might be small, like the Myc and the Flag tails, or intact enzymes, like glucose oxidase, or both.

From all the above described constructs, derived from pUR4421, an appropriate Eagl-HindIII fragment, encoding the functionalized V_H fragment, can be isolated and cloned into a number of different expression plasmids. Several are exemplified in more detail in the following Examples. Although only the V_H fragments are exemplified, similar constructs can be prepared for the production of larger heavy chain fragments (e.g. V_H-hinge or V_H-hinge-C_H2) or intact heavy chains. The Eagl site is introduced before the first codon of the V_H fragment, facilitating an in frame fusion with different yeast signal sequences.

In particular cases, were additional Eagl and/or HindIII sites are present in the cloned fragments, it is necessary to perform partial digestions with one or both restriction enzymes.

Although the above and following constructions only consider the V_{II} fragment cloned in pB3, a comparable construction route can be used for the construction of expression plasmids for the production of V_{II} fragments like V_H-09 and V_H-24, or other V_{II} fragments.

Example 2 Construction of S. cerevisiae episomal expression plasmids for Camelidae V_{II}.

For the secretion of recombinant protein from *S. cerevisiae* it is worthwhile to test in parallel the two most frequently applied homologous signal sequences, the SUC2 invertase signal sequence and the prepro- α mating factor sequence.

- The episomal plasmid pSY1 and pSY16 (Harmsen et al., 1993) contain expression cassettes for the α-galactosidase gene. Both plasmids contain the GAL7 promoter and PGK terminator sequences. pSY1 contains the invertase (SUC2) signal
 - sequence and pSY16 contains a slightly modified (Harmsen et al., 1993) prepro-α-
- 10 mating factor signal sequence.
 - Both plasmids, pSY1 and pSY16 can be digested with *EagI* and *HindIII*, the about 6500 bp long vector backbone of both plasmids can be isolated and subsequently ligated with the *EagI/HindIII* fragments from pUR4421-03F (~465 bp), pUR4421-03M (~455 bp) or pUR4421-03 (~405 bp) (See above).
- This results in a series of 6 different episomal plasmids for expression in S. cerevisiae, containing behind the SUC2- and the α mating factor prepro-sequence the V_H-Flag coding sequence (designated pUR4423F and pUR4426F), the V_H-Myc coding sequence (designated pUR4423M and pUR4426M) or the coding sequence of V_H followed by a stop codon (designated pUR4423, Figure 7 and pUR4426,
- 20 Figure 8).

Obviously, it is possible to use promoter systems different from the inducible GAL7 promoter, e.g. the constitutive GAPDH promoter.

2.1 Production of V_{II} -03-myc and V_{II} -24-myc.

- After introducing the expression plasmids pUR4423M (coding for V_H-03-myc, preceded by the SUC2-signal sequence) and pUR4425M (coding for V_H-24-myc. preceded by the SUC2-signal sequence) into *S. cerevisiae* via electroporation, transformants were selected from minimal medium agar plates (comprising 0.7 % yeast nitrogen base, 2 % glucose and 2 % agar, supplemented with the essential amino acids and bases).
 - For the production of antibody fragments the transformants were grown overnight in selective minimal medium (comprising 0.7 % yeast nitrogen base, 2 % glucose,

supplemented with the essential amino acids and bases) and subsequently diluted ten times in YPGal medium (comprising 1 % yeast extract, 2 % bacto pepton and 5 % galactose). After 24 and 48 hours of growth, samples were taken for Western blot analysis (Figure 21). For the immuno detection of the produced V_{II} -myc fragments monoclonal anti-myc antibodies were used.

In essentially the same way comparable results were obtained with a yeast transformed with pUR4424M containing a DNA sequence encoding the $V_{\rm H}$ -09-myc protein.

Example 3 Construction of *S. cerevisiae* multicopy integration vectors for the expression of *Camelidae* V_{II}.

To combine the benefits of high copy number and mitotically stable expression, the concept of a multicopy integration system into the rDNA locus of lower eukaryotes has already been successfully proven (Giuseppin et al. supra).

One of these vectors is pUR2778, a derivative of pUR2774 (Giuseppin et al. supra) from which the pol1-S.O. reporter gene sequence was removed (Figure 9).

This integrating plasmid, pUR2778, can be used for integration of *Camelidae* V_H coding sequences, hence the vector can be digested with *SacI* and *HindIII* after which the ~7.3 kb vector fragment can be isolated.

From the in example 2 described pUR4423 or pUR4426 types of plasmids, SacI20 HindIII fragments can be isolated encoding a V_H fragment preceded by a signal sequence (SUC2 or α mating factor prepro) and followed by nothing or a Myc or Flag tail.

Ligation of these SacI-HindIII fragments with the ~7.3 kb vector fragment will result in integration plasmids, encoding the (functionalized) V_H fragments under the regulation of the strong and inducible GAL7 promoter.

In this way the following expression plasmids were obtained:

20

pUR4429 P_{gal7} - SUC2 sig.seq. - V_{II} -03 pUR4429F P_{gal7} - SUC2 sig.seq. - V_{II} -03 - Flag tail pUR4429M P_{gal7} - SUC2 sig.seq. - V_{II} -03 - Myc tail pUR4430 P_{gal7} - α mat.fac. prepro. - V_{II} -03 - Flag tail pUR4430F P_{gal7} - α mat.fac. prepro. - V_{II} -03 - Flag tail pUR4430M P_{gal7} - α mat.fac. prepro. - V_{II} -03 - Myc tail

For schematic drawings see Figure 10 for pUR4429 and Figure 11 for pUR4430. Obviously, comparable constructs can be prepared for other heavy chain antibodies or fragments thereof.

As mentioned before, different promoters might be used, for example, the constitutive GAPDH promoter.

Example 4 Construction of expression plasmids for the production of (functionalized) V_{11} fragments from Camelidae by Kluyveromyces

4.1. Construction of *Kluyveromyces lactis* episomal expression plasmids *Camelidae*.

Yeast strains of the genus *Kluyveromyces* have been used for the production of enzymes, such as \(\beta\)-galactosidase for many years, and the growth of the strains has been extensively studied. *Kluyveromyces lactis* is well known for the ability to utilize a large variety of compounds as carbon and energy sources for growth. Since these strains are able to grow at high temperatures and exhibit high growth rates, they are promising hosts for industrial production of heterologous proteins (Hollenberg, C. *et al.*, EP-A-0096430, GIST-BROCADES N.V., 1983).

The plasmids pUR2427 and pUR2428 are pTZ19R derivatives with the promoter and the DNA sequence encoding either the signal peptide (=pre-sequence) (in pUR2428), or the natural prepro-sequence (in pUR2427), of inulinase (inu) from Kluyveromyces marxianus. Both plasmids contain a unique BspMI site suitable to create a perfect joint with EagI or Not1 digested DNA-fragments (not yet published European patent application 92203932.6, supra). In both plasmids a unique HindIII site is located a bit further downstream of the BspMI-site, so that EagI-HindIII cut DNA-fragments encoding V_{II} from Camelidae either solely or with Myc- or Flag- tail

can be easily ligated into *BspMI-HindIII* digested pUR2427 or pUR2428. Thereby a set of six plasmids can be created containing the promoter and secretion signals of the *Kluyveromyces marxianus* inulinase gene, joint in frame to *Camelidae* Vh encoding sequences, all on a *Eco*RI-*HindIII* restriction fragment:

5 pUR4445 P_{inu} - Inu prepro seq. - V_{II} - 03
pUR4445M P_{inu} - Inu prepro seq. - V_{II} - 03 - Myc
pUR4445F P_{inu} - Inu prepro seq. - V_{II} - 03 - Flag
pUR4446 P_{inu} - Inu pre seq. - V_{II} - 03
pUR4446M P_{inu} - Inu pre seq. - V_{II} - 03 - Myc
10 pUR4446F P_{inu} - Inu pre seq. - V_{II} - 03 - Flag .

Maps of pUR4445 and pUR4446 are shown in Figure 12 and Figure 13.

The EcoRI-HindIII fragments of these plasmids can be ligated into the expression vector pSK1 (not yet published European patent application 92203932.6, supra),

from which the α-galactosidase expression cassette including the GAL7-promoter is removed with a EcoRI(partial) and HindIII digestion. The resulting plasmids can then be transformed for example in K. lactis strain MSK110 (a, uraA, trp1::URA3), as they contain the trp1 marker and the pKD1 episomal plasmid sequences:

pUR4447 P_{inu} - Inu prepro seq. - V_H - 03 20 pUR4447M P_{inu} - Inu prepro seq. - V_H - 03 - Myc pUR4447F P_{inu} - Inu prepro seq. - V_H - 03 - Flag pUR4448 P_{inu} - Inu pre seq. - V_H - 03 pUR4448M P_{inu} - Inu pre seq. - V_H - 03 - Myc pUR4448F P_{inu} - Inu pre seq. - V_H - 03 - Flag .

25 A map of pUR4447 is shown in Figure 14.

Transformation can be performed by standard techniques such as the methods of Beggs (1978) or electroporation, using 0.67% Yeast Nitrogen Base (without amino acids) and 2% glucose as the selection medium for transformants.

4.2. Construction of Kluyveromyces lactis multicopy integration vectors.

Alternatively, since all tailed and non-tailed versions of the Vh fragments, joined to the inulinase promoter and secretion signals, are located on EcoRI-HindIII fragments, the rDNA multicopy integration plasmid pMIRKGAL-Tal (Bergkamp et al., 1992) can be used in a similar way as the pSK1 plasmid. In order to replace the α-gal expression cassette present in this plasmid, by a antibody fragment cassette, these plasmids have to be digested with EcoRI(partial) and HindIII. After isolating the vector fragments, they can be ligated with the about 1.2 kb EcoRI-HindIII fragments which can be obtained from the plasmids described in example 4.1. The resulting plasmids can be linearized with SacII and transformed to MSK110, resulting in K. lactis strains with potentially high and stable expression of single chain V_H fragments.

pUR4449 P_{ing} - Inu prepro seq. - V_{II} - 03 P_{inu} - Inu prepro seq. - V_H - 03 - Myc pUR4449M 15 pUR4449F P_{inu} - Inu prepro seq. - V_H - 03 - Flag pUR4450 P_{inu} - Inu pre seq. - V_H - 03 P_{inu} - Inu pre seq. - V_H - 03 - Myc pUR4450M P_{inu} - Inu pre seq. - V_H - 03 - Flag. pUR4450F

10

20 4.3. Construction of Kluyveromyces marxianus episomal plasmids.

Kluyveromyces marxianus is a yeast which is perhaps even more attractive than K. lactis for industrial biotechnology, due to its short generation time on glucose (about 45 minutes) and its ability to grow on a wide range of substrates, and its growth at elevated temperatures (Rouwenhorst et al., 1988).

25 The shuttle vector pUR2434, containing the leu2 marker and the pKD1 plasmid sequences (not yet published European patent application 92203932.6, supra), located on a pUC19 based vector, can be cut with EcoRI(partial) and HindIII, to remove the α-galactosidase expression cassette. In this vector the EcoRI-HindIII fragments containing the Vh expression cassettes as described in example 4.1, can be 30 ligated. The resulting plasmids can then be transformed into KMS3, the neat leu2auxotroph CBS6556 K. marxianus strain (Bergkamp, 1993) using the method of Meilhoc et al. (1990).

WO 94/25591 PCT/EP94/01442

```
pUR4451 P_{inu} - Inu prepro seq. - V_{II} - 03

pUR4451M P_{inu} - Inu prepro seq. - V_{II} - 03 - Myc

pUR4451F P_{inu} - Inu prepro seq. - V_{II} - 03 - Flag

pUR4452 P_{inu} - Inu pre seq. - V_{II} - 03

pUR4452M P_{inu} - Inu pre seq. - V_{II} - 03 - Myc

pUR4452F P_{inu} - Inu pre seq. - V_{II} - 03 - Flag .

A map of pUR4451 is shown in Figure 15.
```

4.4 Construction of *Kluyveromyces marxianus* multicopy integration vectors.

10 For high and stable expression in *Kluyveromyces marxianus*, the multicopy integration system as described by Bergkamp (1993), can be used. The following cloning route, based on the route for constructing pMIRKM-GAL5 (Bergkamp, 1993), results in suitable expression vectors for production of Vh fragments from Camelidae. The EcoRI-NheI(Klenow filled) fragments of pUR4447,-M,-F and pUR4448,-M,-F 15 containing the Vh fragment expression cassettes as described in example 4.1, can be isolated and ligated in EcoRI-EcoRV digested pIC-20H. From the plasmids obtained in this way, and which are equivalents of the pIC-agal plasmid, the BamHI-NruI fragment can be isolated and ligated with BamHI-SmaI digested pMIRKM4. The result of this will be expression vectors which are equivalent to pMIRKM-GAL5, 20 and contain a tailed or non-tailed Vh fragment from camel under control of inulinase promoter and secretion signals, in a vector which also contains the K. marxianus LEU2-gene with defective promoter, and K. marxianus rDNA sequences for targeted integration into the genome. These vectors can be used to transform for example KMS3.

15

20

25

Example 5. Construction of Hansenula polymorpha integrating vectors for the expression of (functionalized) V_H fragments from Camelidae.

In search for productive systems able to carry out authentic posttranscriptional processing and overcoming the limitation of higher eukaryotic expression systems, such as high costs, low productivity and the need for stringent control procedures for the detection of contaminating agents could be overcome by the methylotrophic yeast *H. polymorpha*. This strain is able to grow on methanol as its sole carbon and energy source, so the presence of methanol in the growth medium rapidly induces the enzymes of the methanol pathway, such as the key enzymes methanol oxidase (MOX) and dihydroxyacetone synthase (DHAS).

While experiments to express foreign genetic information from an episomal plasmid resulted a low plasmid stability, chromosomal integration is the method of choice (Sierkstra et al., 1991). By utilizing the DNA of the mox gene as integration locus the latter were able to express and secrete α -galactosidase regulated by mox promoter and -terminator. Here, the S. cerevisiae SUC2 signal sequence was proven to be efficiently functional for secretion.

The same approach can be used for expression and secretion of *Camelidae* V_H antibody fragments. Plasmids analogous to pUR3515 (without an origin of replication functional in yeast) and pUR3517 (containing the HARS2 sequence as origin of replication) can be used as expression vectors (Sierkstra *et al.*, 1991). As a starting vector pUR3501 can be used (Sierkstra *et al.*, 1991) in which by means of site directed mutagenesis (e.g. via PCR technology), an *Eag*I restriction site is introduced at the junction between the invertase (=SUC2) signal sequence and the α-galactosidase. From the resulting plasmid, pUR3501*Eag*, it is possible to replace the *Eag*I-*Hin*dIII fragment comprising the α-galactosidase gene by an *Eag*I-*Hin*dIII fragment encoding a (functionalized) antibody fragment, obtained as described in example 1. In case of using the *Eag*I-*Hin*dIII fragments of the pUR4421-03 series (example 1), this would result in plasmids pUR4437 (Figure 17), pUR4437M and pUR4437F. In these plasmids the nucleotide sequence encoding the (functionalized) V_{II} is preceded by a nucleotide sequence encoding the invertase signal sequence and the *mox* promoter sequence. The obtained plasmids can be digested with *Bam*HI

and HindIII and after filling in the sticky ends with Klenow polymerase, the about

2.6 kb fragments can be ligated into plasmid pUR3511 which was digested with SmaI (Sierkstra et al., 1991). In this way the terminator sequence of the mox gene can by fused downstream of the V_{II} encoding sequences. From the thus obtained plasmids, pUR4438 (Figure 17) EcoRI-HindIII fragments of about 3 kb can be isolated, containing the mox promoter, the invertase signal sequence, the (functionalized) V_{II} fragment and the mox transcription terminator. Subsequently these fragments can be cloned into plasmid pUR3513 (no yeast origin of replication) or in pUR3514 (HARS origin of replication) as described by Sierkstra et al. (1991), resulting in two sets of plasmids:

10

```
pUR4439 P_{mox} - SUC2 sig. seq. - V_H - mox term. -- no origin pUR4439M P_{mox} - SUC2 sig. seq. - V_H - mox term. -- no origin pUR4439F P_{mox} - SUC2 sig. seq. - V_H - mox term. -- no origin pUR4440 P_{mox} - SUC2 sig. seq. - V_H - mox term. -- HARS origin pUR4440M P_{mox} - SUC2 sig. seq. - V_H - mox term. -- HARS origin pUR4440F P_{mox} - SUC2 sig. seq. - V_H - mox term. -- HARS origin . Maps of pUR4439 and pUR4440 are shown in Figure 18.
```

Essentially the same can be done with other *EagI-HindIII* fragment, obtained as described in example 1.

The newly obtained plasmids can be transformed by electroporation of *H.* polymorpha A16 (CBS4732, leu-) and can be selected by growing on selective medium containing 0.68% YNB and 2% glucose. Induction medium should contain 0.5% methanol instead of the glucose.

25

20

Example 6 Construction Aspergillus niger var. awamori integration vectors for the production of V_{II} fragments from Camelidae.

The multiple cloning site of plasmid pEMBL9 (ranging from the *Eco*RI to the *Hind*III site) was replaced by a synthetic DNA fragment having the nucleotide sequence as indicated in Figure 19; see SEQ. ID. NO: 42-45. The 5'- part of the nucleotide sequence contains a *Nru*I restriction site followed by the first codons of the *Camelidae* V_{II} gene fragment and a *Xho*I restriction site. The 3'-part encodes for

a BstEII restriction site, the last codons of the $Camelidae\ V_{II}$ gene, eleven codons of the Myc tail and finally a EcoR1 and a AfIII site. The resulting plasmid is pUR4432.

After digesting plasmid pB3 with Xhol and EcoRI, a DNA fragment of approximately 425 bp can be isolated from agarose gel. This fragment codes for a truncated V_{II}-Flag fragment, missing the first 5 amino acids of the Camelidae V_{II}.

The obtained fragment can be cloned into pUR4432. To this end plasmid pUR4432 can be digested with Xhol and EcoRI, after which the about 4 kb vector fragment was isolated from an agarose gel. Ligation with the about 425 bp fragment resulted in plasmid pUR4433F.

After digesting the plamids pB3 with XhoI and BstEII, a DNA fragment of approximately 365 bp was isolated from agarose gel. This fragment codes for a truncated $V_{\rm H}$ fragments, missing the first and last 5 amino acids of the Camelidae $V_{\rm H}$.

The obtained fragment was cloned into pUR4432. To this end plasmids pUR4432 can be digested with Xhol and BstEII, after which the about 4 kb vector fragment was isolated from an agarose gel. Ligation with the about 365 bp fragments resulted in plasmids pUR4433M. In a similar way the Xhol-BstEII fragments of pB9 and pB24 were cloned into the pUR4432 vector fragment, resulting in pUR4434M and pUR4435M, respectively.

Upon digesting pUR4433M or pUR4433F with *BstEII* and *HindIII*, the vector fragments of about 4.4 kb can be isolated from agarose gel and religated in the presence of a synthetic linker peptide having the following sequence:

BSTEII AflII HindIII
25 GTCACCGTCTCCTCATAATGATCTTAAGGTGATA
GCAGAGGAGTATTACTAGAATTCCACTATTCGA (see SEQ. ID. NO: 50-51).

In the thus obtained plasmid, pUR4433, the Myc tail or Flag tail sequences are removed and the $V_{\rm H}$ gene fragment is directly followed by a stop codon.

Analogous as described in example 1.5, it is possible to clone nucleotide sequences encoding longer fragments of the heavy chain immunoglobulins into pUR4432 or to replace the *BstEII-AflII* fragments of the above mentioned plasmids pUR4433,

WO 94/25591 PCT/EP94/01442

pUR4433F or pUR4433M with other BstEII-AfIII fragments, resulting in frame fusions encoding functionalized V_{II} fragments, having a C-terminal extension. Upon replacing the NruI-XhoI fragments of pUR4433, pUR4433F or pUR4433M, in frame fusions can be constructed encoding functionalized V_{II} fragments, having an

5 N-terminal extension.

In the above described constructs an Nrul site was introduced before the first codon of the (functionalized) V_{II} fragment, facilitating an in frame fusion with the precursor-sequence of xylanase, see (not prior-published) WO-A-93/12237, supra. For the construction of Aspergillus expression plasmids, from the plasmids pUR4433F, pUR4433M and pUR4433, respectively, an about 455, 445 and 405 bp Nrul-AflII fragment has to be isolated encoding the V_{II} fragment with a Flag, a Myc or no tail.

Plasmid pAW14B was the starting vector for construction of a series of expression plasmids containing the exlA expression signals and the genes coding for (functionalized) V_H fragments of Camelidae heavy chain antibodies. The plasmid comprises an Aspergillus niger var. awamori chromosomal 5 kb SalI fragment on which the 0.7 kb exlA gene is located, together with 2.5 kb of 5'-flanking sequences and 2.0 kb of 3'-flanking sequences (see Figure 20 and (not prior-published) WO-A-93/12237, supra).

Starting from pAW14B, pAW14B-10 was constructed by removing the *Eco*RI site originating from the pUC19 polylinker, and introducing a *Not*I site. This was achieved by digesting plasmid pAW14B with *Eco*RI and after dephosphorylation the linear 7.9 kb *Eco*RI fragment was isolated. The fragment was religated in the presence of the "*Eco*RI"-*Not*I linker:

5'- AATTGCGGCCGC -3'

(see SEQ. ID. NO: 52).

Subsequently the AfIII site, located downstream of the exlA terminator was removed by partially cleaving plasmid pAW14B-10 and religating the isolated, linearized plasmid after filling in the sticky ends, resulting in plasmid

30 pAW14B-11.

25

Finally, pAW14B-12 was constructed using pAW14B-11 as starting material. After digestion of pAW14B-11 with AfIII (overlapping with the exlA stop codon) and BglII

(located in the *exl* promoter) the ~2.4 kb *AfIII-BgIII* fragment, containing part of the *exlA* promoter and the *exlA* gene was isolated as well as the ~5.5 kb *AfIII-BgIII* vector fragment. After partial digestion of this ~2.4 kb fragment with *BspHI* (located in the *exlA* promoter and at the *exlA* start codon) an about 1.8 kb *BgIII-BspHI exlA* promoter fragment (up to the ATG initiation codon) was isolated and ligated with the about 5.5 kb *AfIII-BgIII* vector fragment of pAW14B-11 in the presence of the following adaptor:

(BspHI) BbsI AflII
CATGCAGTCTTCGGGC

(see SEQ. ID. NO: 53-54).

10 GTCAGAAGCCCGAATT

presence of the antibody fragments.

For the construction of the V_{II} expression plasmids, pAW14B-11 can be partially digested with NruI and digested with AfIII, after which the ⁻ 7 kb vector fragment can be isolated from agarose gel and contains the xylanase promoter, the DNA sequence encoding the xylanase signal sequence and the xylanase terminator. Upon ligation of the NruI-AfIII fragments of pUR4433M, pUR4434M and pUR4435M with the pAW14B-11 vector, plasmids pUR4436M, pUR4437M and pUR4438M were obtained, respectively. In these plasmids the Camelidae V_H polypeptides are preceded by the 27 amino acid long precursor sequence of xylanase and followed by the myc-tail (of 11 amino acids; see Examples 1.3 en 2, Figures 6 and 19, and SEQ.ID. NO: 41 = 45).

In a similar way plasmids can be constructed encoding the V_H fragments followed by the FLAG-tail or without a tail.

After introducing the amdS and pyrG selection markers into the unique NotI site of pUR4436M, pUR4437M and pUR4438M using conventional techniques, e.g. as

described in Examples 2 and 3 of (not prior-published) WO-A-93/12237, supra, the plasmids were transferred to Aspergillus.

Production of the Camel V_{II} fragments by the selected transformants was achieved by growing the strains in inducing medium essentially as described in example 2,2 of (not prior-published) WO-A-93/12237, *supra*. Western blot analysis of the culture medium was perforemed as described in Example 2.1 above and revealed the

Obviously, expression vectors can be constructed in which different promoter systems, e.g. glucoamylase promoter, and/or different signal sequences, e.g. glucoamylase or glucose oxidase signal sequences, are used.

5 Example 7 Production of glucose oxidase - V_{II} fusion proteins

Glucose oxidase catalyses the oxidation of D-glucose to D-gluconate under the release of hydrogen peroxide. Glucose oxidase genes (gox) from Aspergillus niger have been cloned (Frederick et al. (1990) J. Biol. Chem. 265 3793, Kriechbaum et al., 1989) and the nucleotide sequences are available from the EMBL data bank under accession numbers J05242 and X16061. The nucleotide sequence of the latter is used as a basis for the following construction route.

Upon cloning the gox gene from A. niger it is possible, by applying PCR technology, to introduce convenient restriction sites.

To introduce a *Bsp*HI restriction site, overlapping with the ATG initiation codon,
the sequence ATC ATG CAG can be changed to ATC ATG AGG. In the same
experiment an *Eco*RI restriction site can be introduced which is located upstream of
the *Bsp*HI site. This can be achieved by using the following PCR primer:

ECORI BspHI
5'-TCACTGAATTCGGGATC ATG AGG ACT CTC CTT GTG AGC TCG CTT-3'
(see SEO, ID, NO: 55).

A second PCR primer, having the following sequence can be used:

AflII BbsI SalI
5'-ATGTCACAAAGCTTAAGCACGAAGACA GTC GAC CGT GCG GCC GGA GAC-3'
HindIII

25 (see SEQ. ID. NO: 56)

20

30

in the same PCR experiment, in order to introduce a *BbsI* site, a *AfIII* site and a *HindIII* site, downstream of the unique *SalI* site present in the glucose oxidase gene. After digesting the DNA obtained from this PCR experiment with *EcoRI* and *HindIII*, an *EcoRI* - *HindIII* fragment of about 160 bp can be isolated and cloned into pEMBL9, which was digested with the same enzymes, resulting in plasmid pGOX1.

From pGOX1 an about 140 bp BspHI - AfIII fragment can be isolated and introduced into the 7.2 kb BbsI-AfIII vector fragment of pAW14B-12, resulting in

15

pAW14B-GOX. In this plasmid, the 5'- part of the gox gene, encoding the first 43 amino acids, is fused in frame with the ATG initiation codon of the exlA gene.

In a second PCR experiment, a MluI restriction site can be introduced near the 3'end of the gox by changing the sequence TAT GCT TCC to TAC GCG TCC. In the same experiment a HindIII site can be introduced downstream of the MluI site. As a second primer an oligo nucleotide should be used hybridizing upstream of the Sall site. After digesting the DNA obtained from this PCR experiment with SalI and HindIII, an Sall - HindIII fragment of about 1.7 kb can be isolated and cloned into pEMBL9, which was digested with the same enzymes, resulting in plasmid pGOX2. Upon digesting pGOX2 with MluI and HindIII, an about 5.7 kb vector fragment can be isolated.

From the plasmids pUR4433, pUR4433F, pUR4433M and the like, XhoI-HindIII fragments can be isolated, encoding the truncated Camelidae V_{II} fragment with or without a tail sequence, and missing the first 4-6 N-terminal amino acids (see Example 1). These fragments can be ligated into the 5.7 kb pGOX2 vector fragment by using MluI-XhoI adaptors. These adaptors are designed in such a way that they result in an in frame fusion between the 3'-end of the gox gene and the restored V_H gene fragment, optionally intersected with a DNA sequence encoding a peptide 20 linker sequence.

An example of these designed adaptors is:

XhoI CGCGTCCATGCAGTCCTCAGGTGGATCATCCCAGGTGAAACTGC AGGTACGTCAGGAGTCCACCTAGTAGGGTCCACTTTGACGAGCT 25 S G G S S | Q V K L L E S M Q | S(see SEQ. ID. NO: 57-59)

which encodes for the last amino acids of GOX, an SSGGSS linker sequence (see SEQ. ID. NO: 62) and the N-terminal amino acids of the Camel V_H fragment of pB3. Instead of the SSGGSS linker (see SEQ. ID. NO: 62) it is possible to use other linkers such as the repeated sequence linkers described in the above indicated European patent application 92402326.0, e.g. a repeated sequence Pro-X, with X being any amino acid, but preferably Gln, Lys or Glu, the sequence containing

WO 94/25591 PCT/EP94/01442

40

advantageously at least 3 repeats of Pro-X and especially a fragment composed of a 12-fold repeat of the sequence Pro-X.

In case the about 435 bp *Xhol-HindIII* fragment of pUR4433M is used in combination with the above described adaptor, this would result in pGOX2-03M. From this plasmid a *Sall-AfIII* fragment of about 2.1 kb encoding the C-terminal part of glucose oxidase followed by the linker peptide, the Camel V_{II} fragment of pB3 and finally the Myc tail.

Upon digesting pAW14B-GOX partially with *Bbs*I, and with *Afl*II, the about 7.4 kb vector fragment can be isolated. This fragment contains the xylanase promoter, the DNA sequence encoding the N-terminal part of glucose oxidase and the xylanase promoter. Due to the digestion with *Bbs*I, a *Sal*I sticky end is created, corresponding with the *Sal*I restriction site originally present in the *gox* gene. Ligation of the *Sal*I-AflII vector fragment with the about 2.1 kb *Sal*I-AflII fragment of pGOX2-03M,

resulting in pUR4441M. This expression plasmid encodes for a single chain polypeptide comprising the glucose oxidase enzyme, the (functionalized) Camel V_H fragment and the Myc tail.

Introduction of this type of expression plasmids in *Aspergillus* can be achieved essentially as described in example 6.

As the naturally occurring glucose oxidase is a homodimeric enzyme, it might be expected that a fusion protein, comprising glucose oxidase and an antibody fragment as a C-terminal extension, has an increased avidity for the antigen/antibody binding, if this fusion protein is produced as a homodimer. Alternatively, it is possible to produce heterodimers, consisting of one glucose oxidase molecule connected to a V_H fragment and one wild type glucose oxidase molecule. This can be achieved by producing with the same strain both wild type glucose oxidase and the fused glucose oxidase-V_H fragment, or by mixing the two different homodimers produced by different strains under conditions whereby the mixture of dimers are dissociated and subsequently associated.

20

Example 8 Engineering of Camelidae V_{II} fragments

8.1 Random and targeted random mutagenesis.

After expressing a number of different Camelidae V_{II} fragments in lower eukaryotic host organisms as described above, or in prokaryotes, fragments produced in relative higher amounts can be selected. Upon subjecting the XhoI-BstEII gene fragments to a (targeted) random mutagenesis procedure, it might be possible to further improve special characteristics of the V_{II} fragment, e.g. further improvement of the production level, increased stability or increased affinity.

To this end the following procedure might be followed.

10 Upon replacing the polylinker of the phagemid vector pHEN1 (Hoogenboom et al., 1991) located on a Ncol-NotI fragment by a new polylinker having the following sequence:

NCOI XhOI BSTEII NOTI
CATGGCCAGGTGAAACTGCTCGAGTAAGTGACTAAGGTCACCGTCTCCTCAGC
CGGTCCACTTTGACGAGCTCATTCACTGATTCCAGTGGCAGAGGAGTCGCCGG

(see SEQ. ID. NO: 60-61) it becomes possible to introduce XhoI-BstEII fragments encoding truncated Camelidae V_H fragments in the phagemid.

- Following mutagenesis of the V_H encoding sequence (random mutagenesis) or a specific part thereof (targeted random mutagenesis), the mutated V_H fragments can be expressed and displayed on the phage surface in essentially the same way as described by Hoogenboom *et al.* (1991). Selecting phages displaying (mutant) V_H fragments, can be done in different ways, a number of which are described by Marks *et al.* (1992). Subsequently, the mutated *XhoI-Bst*EII fragments can be isolated from
- 25 the phagemid and introduced into expression plasmids for yeast or fungi as described in previous examples.
 - Upon producing the mutant V_{II} fragments by these organisms, the effects of the mutations on production levels, V_{II} fragment stability or binding affinity can be evaluated easily and improved V_{II} fragments can be selected.
- 30 Obviously, a similar route can be followed for larger antibody fragments. With similar procedures the activity of catalytic antibodies can be improved.

8.2 Site-directed or designed mutagenesis

As an alternative to the methods described above in Example 8.1 it is possible to use the well-known technique of site-directed mutagenesis. Thus, designed mutations, preferably based on molecular modelling and molecular dynamics, can be introduced in the V_{11} fragments, e.g. in the framework or in the CDRs.

8.3 Construction V_{II} fragments with regulatable binding efficiencies.

For particular applications, the possibility to regulate the binding capacity of antibody fragments might be necessary. The introduction of metal ion binding sites in proteins is known from the literature e.g. Pessi et al. (1993). The present inventors envisage that the introduction of a metal binding site in an antibody fragment by rational design can result in a regulatable antibody fragment, when the metal binding site is introduced at a position such that the actual binding of the metal ion results in a conformational change in the antibody fragments due to which the binding of the antigen to the antibody fragment is influenced. Another possibility is that the presence of the metal prevents antigen binding due to steric hindrance.

8.4 Grafting of CDR regions on the framework fragments of a Camelidae V_H fragment.

- Grafting of CDR fragments onto framework fragments of different antibodies or fragments thereof is known from the literature (see Jones et al. (1986), WO-A-92/15683, and WO-A-92/01059). In these cases the CDR fragments of murine antibody fragments were grafted onto framework fragments of human antibodies. The sole rationale behind the "humanization" was to increase the acceptability for therapeutic and/or diagnostic applications in human.
 - Essentially the same approach can however also be used for a totally different purpose. Although antibody fragments share some homology in the framework areas, the production levels vary considerably.
- Once an antibody or an antibody fragment, e.g. a *Camelidae* V_{II} fragment, has been identified, which can be produced to high levels by an production organism of interest, this antibody (fragment) can be used as a starting point to construct "grafted" antibody (fragments), which can be produced in high levels and have an

WO 94/25591 PCT/EP94/01442

other specificity as compared to the original antibody (fragment). In particular cases it might be necessary to introduce some modifications in the framework fragments as well in order to obtain optimal transitions between the framework fragments and the CDR fragments. For the determination of the optimal transitions molecular

5 dynamics and molecular modelling can be used.

To this end a synthetic gene, encoding the "grafted V_H " fragment, can be constructed and introduced into an expression plasmid. Obviously it is possible to adapt the codon usage to the codons preferred by the host organism.

For optimization of the "grafted V_{II} " fragment, the procedure as described in example 8.1 can be followed.

Literature mentioned in the specification additional to that mentioned in the above given draft publication

- Adair, J.R. et al., WO-A-92/01059 (CELLTECH Ltd, 1992)
- 15 Beggs (1978) Nature 275 104
 - Bendig, M.M. et al. WO-A-92/15683 (MERCK PATENT GmbH, 1992)
 - Bergkamp, R.J.M., Kool, I.M., Geerse, R.H., Planta, R.J. (1992) Multiple copy integration of the α-galactosidase gene from *Cyamopsis tetragonoloba* into the ribosomal DNA of *Kluyveromyces lactis*. Current Genetics 21 365-370
- Bergkamp, R.J.M., PhD Thesis Free University of Amsterdam (1993),
 Heterologous gene expression in Kluyveromyces yeasts
 - Better et al. (1988) Science 240 1041-1043
 - Bird et al., (1988) Science 242 423-426
 - Cabilly, S. et al., EP-A-0125023 (GENENTECH, 1984)
- 25 Denthe, et al. (1983) Nucl. Acids Res. 11 1645
 - Fellinger, A.J. et al., EP-A-0255153 (UNILEVER, 1988)
 - Frederick et al. (1990) J. Biol. Chem. 265 3793
 - Giuseppin, M.L.F., Lopes, M.T.S., Planta, R.J., Verbakel, J.M.A., Verrips, C.T. (1991) Process for preparing a protein by a yeast transformed by multicopy
- integration of an expression vector. PCT application WO 91/00920 (UNILEVER)
 - Harmsen, M.M., Langedijk, A.C., van Tuinen, E., Geerse, R.H., Rauè, H.A., Maat, J., (1993) Effect of pmr1 disruption and different signal sequences on the

- intracellular processing and secretion of *Cyamopsis tetragonoloba* α -galactosidase by *S. cerevisiae*. Gene 125 115-123
- Hollenberg, C. et al., EP-A-0096430 (GIST-BROCADES N.V., 1983))
- Hoogenboom H.R., Griffiths, A.D., Johnson, K.S., Chiswell, D.J., Hudson, P., and
- Winter, G. (1991) Multi-subunit proteins on the surface of filamentous phage: methodologies for displaying antibody (Fab) heavy and light chains. Nucleic Acids Research 15 4133-4137
 - Jones et al. (1986) Nature 321 522
 - Kriechbaum et al. (1989) FEBS Lett. 255 63
- 10 Ledeboer, A.M. *et al.*, EP-A-0173378 (UNILEVER, 1986)
 - Leenhouts, C.J. et al., EP-A-0487159 (UNILEVER, 1992)
 - Lerner, Benkovic and Schultz, (1991) Science 252 659-667
 - Marks, J.D., Hoogenboom, H.R., Griffiths, A.D., and Winter, G. (1992) Molecular evolution of proteins on filamentous phage. J. Biol. Chem. 267 16007-16010
- Meilhoc, E., Masson, J., Teissié, J. (1990) High efficiency transformation of intact yeast cells by electric pulses. Bio/Technology 8 223-227
 - Mitchell, P., (1979) Science 206 1148-1159)
 - Pessi et al. (1993) Nature 362 367.
 - Rouwenhorst, R.J., Visser, L.E., van der Baan, Scheffers, W.A., van Dijken, J.P.
- 20 (1988) Production, distribution and kinetic properties of inulinase in continuous culture of *Kluyveromyces marxianus* CBS 6556. Appl. Environm. Microbiol. 54 1131-1137.
 - Sierkstra, L.N., Verbakel, J.M.A. and Verrips, C.T. (1991) Optimisation of a host/vector system for heterologous gene expression by *Hansenula polymorpha*.
- 25 Current Genetics 19 81-87.
 - Skerra et al. (1988) Science 240 1938
 - Takahashi et al. (1993) Science 259 1460-1463);
 - Teeri et al., WO-A-93/02198 (TECH. RES. CENT. FINLAND, publ. 04.02.1993)
 - Van Gorcom, R.F.M. et al., WO-A-91/19782 (UNILEVER, 1991)
- 30 Wu et al. (1993) Bio/Technology 11 71
 - Zhou et al. (1991) Nucleic Acids Research 19 6052

Additional references to prior-filed but not prior-published patent applications, which are incorporated herein by reference:

- not prior-published PCT application EP 92/02896, filed 09.12.92 with priority date of 09.12.91 (UNILEVER / R.J. Gouka *et al.*), now publicly available as WO-A-93/12237
- not prior-published EP application 92202080.5, filed <u>08.07.92</u> (UNILEVER / F.M. Klis *et al.*), now publicly available as International (PCT) patent application WO-A-94/01567)
- not prior-published EP application 92402326.0, filed <u>21.08.92</u> (C. Casterman & R. Hamers), now publicly available as EP-A1-0 584 421
- not yet published EP application 92203932.6, filed 11.12.92 (UNILEVER / H.Y. Toschka & J.M.A. Verbakel).

15

10

5

Information on deposits of micro-organisms under the Budapest Treaty is given in Example 1 on page 23, lines 23-25 above. In agreement with Rule 28 (4) EPC, or a similar arrangement for a State not being a Contracting State of the EPC, it is hereby requested that a sample of such deposit, when requested, will be submitted to an expert only.

SEQUENCE LISTING

(1) GENERAL INFORMATION: 5 (i) APPLICANT: (A) NAME: Unilever N.V. (B) STREET: Weena 455 (C) CITY: Rotterdam (E) COUNTRY: The Netherlands 10 (F) POSTAL CODE (ZIP): NL-3013 AL (A) NAME: Unilever PLC (B) STREET: Unilever House Blackfriars (C) CITY: London 15 (E) COUNTRY: United Kingdom (F) POSTAL CODE (ZIP): EC4P 4BQ (A) NAME: Leon Gerardus Joseph FRENKEN (B) STREET: Geldersestraat 90 20 (C) CITY: Rotterdam (E) COUNTRY: The Netherlands (F) POSTAL CODE (ZIP): NL-3011 MP (A) NAME: Cornelis Theodorus VERRIPS 25 (B) STREET: Hagedoorn 18 (C) CITY: Maassluis (E) COUNTRY: The Netherlands (F) POSTAL CODE (ZIP): NL-3142 KB 30 (A) NAME: Raymond HAMERS (B) STREET: Vijversweg 15 (C) CITY: Sint-Genesius-Rode (E) COUNTRY: Belgium (F) POSTAL CODE (ZIP): B-1640 35 (A) NAME: Cécile HAMERS-CASTERMAN (B) STREET: Vijversweg 15 (C) CITY: Sint-Genesius-Rode (E) COUNTRY: Belgium 40 (F) POSTAL CODE (ZIP): B-1640 (A) NAME: Serge Victor Marie MUYLDERMANS (B) STREET: Brusselse Steenweg 55 (C) CITY: Hoeilaart 45 (E) COUNTRY: Belgium (F) POSTAL CODE (ZIP): B-1560 (ii) TITLE OF INVENTION: Production of antibodies or (functionalized) fragments thereof derived from heavy chain immunoglobulins 50 of Camelidae. (iii) NUMBER OF SEQUENCES: 62 (iv) COMPUTER READABLE FORM: 55 (A) MEDIUM TYPE: Floppy disk (B) COMPUTER: IBM PC compatible (C) OPERATING SYSTEM: PC-DOS/MS-DOS (D) SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO) 60 (2) INFORMATION FOR SEQ ID NO: 1: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 5 amino acids 65 (B) TYPE: amino acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear

```
(ii) MOLECULE TYPE: protein
         (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
5
         Ala Pro Glu Leu Leu
    (2) INFORMATION FOR SEQ ID NO: 2:
10
          (i) SEQUENCE CHARACTERISTICS:
               (A) LENGTH: 5 amino acids
               (B) TYPE: amino acid
               (C) STRANDEDNESS: single
15
               (D) TOPOLOGY: linear
         (ii) MOLECULE TYPE: protein
         (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
20
         Ala Pro Glu Leu Pro
25
    (2) INFORMATION FOR SEQ ID NO: 3:
          (i) SEQUENCE CHARACTERISTICS:
               (A) LENGTH: 21 base pairs
               (B) TYPE: nucleic acid
30
               (C) STRANDEDNESS: single (D) TOPOLOGY: linear
         (ii) MOLECULE TYPE: DNA (genomic)
35
         (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
                                                                                21
    CGCCATCAAG GTACCAGTTG A
40
    (2) INFORMATION FOR SEQ ID NO: 4:
          (i) SEQUENCE CHARACTERISTICS:
               (A) LENGTH: 89 amino acids (B) TYPE: amino acid
45
               (C) STRANDEDNESS: single
               (D) TOPOLOGY: linear
         (ii) MOLECULE TYPE: protein
50
        (vii) IMMEDIATE SOURCE:
               (B) CLONE: human heavy chain framework (subgroup III)
                           (Xaa = CDR1, Xaa Xaa = CDR2 and Xaa Xaa Xaa = CDR3)
55
         (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
         Glu Val Gln Leu Val Glu Ser Gly Gly Leu Val Gln Pro Gly Gly
60
          Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Xaa Trp Val Arg Gln Ala
          Pro Gly Lys Gly Leu Glu Trp Val Ser Xaa Xaa Arg Phe Thr Ile Ser
65
          Arg Asp Asn Ser Lys Asn Thr Leu Tyr Leu Gln Met Asn Ser Leu Arg
              50
                                   55
```

	Ala 65	Glu	Asp	Thr	Ala	Val 70	Tyr	Tyr	Cys	Ala	Arg 75	Xaa	Xaa	Xaa	Trp	. Gl y 80
5	Gln	Gly	Thr	Leu	Val 85	Thr	Val	Ser	Ser							
	(2) INFO	RMATI	ION I	FOR S	SEQ :	ID NO	D: 5	:								
10 15	(i)	(A) (B) (C)	LEI TYI STI	E CHANGTH: PE: 8 RANDI	: 81 amino EDNES	amin cac: SS: 9	no ad id sing:	cids								
13	(ii)	MOLE	CULI	E TYI	PE: 1	prote	∍in									
20	(vii)			TE SO ONE:	came	el "l										k A = CDR3)
	(xi)	SEQ	JENCI	E DES	SCRII	OIT	N: SI	EQ II	ONO:	: 5:						
25	Gly 1	Gly	Ser	Val	Gln 5	Gly	Gly	Gly	Ser	Leu 10	Arg	Leu	Ser	Cys	Ala 15	Ile
	Ser	Gly	Xaa	Trp 20	Phe	Arg	Glu	Gly	Pro 25	Gly	Lys	Glu	Arg	Glu 30	Gly	Ile
30	Ala	Xaa	Xaa 35	Arg	Phe	Thr	Ile	ser 40	Gln	Asp	Ser	Thr	Leu 45	Lys	Thr	Met
35	Tyr	Leu 50	Leu	Met	Asn	Asn	Leu 55	Lys	Pro	Glu	Asp	Thr 60	Gly	Thr	Tyr	Tyr
	Сув 65	Ala	Ala	Xaa	Xaa	Xaa 70	Trp	Gly	Gln	Gly	Thr 75	Gln	Val	Thr	Val	Ser 80
4 0	Ser															
	(2) INFO	RMAT	ON I	FOR S	SEQ :	ID NO	D: 6:	:								
45	(i)	(A) (B) (C)	LEI TYI STI	E CHANGTH: PE: 8 RANDI POLOG	: 81 amino EDNE:	amin ac: SS: :	no ad id sing:	cids								
50	(ii)	MOLE	CULI	E TYI	PE: p	prote	∍in									
55	(Vii)			TE SO ONE:	came	el "h										k B = CDR3)
	(xi)	SEQU	JENCI	E DES	SCRII	PTIO	N: SI	EQ II	ONO:	: 6:						
60	Gly 1	Gly	Ser	Val	Gln 5	Ala	Gly	Gly	Ser	Leu 10	Arg	Leu	Ser	Cys	Ala 15	Ser
•	Ser	Ser	Xaa	Trp 20	Tyr	Arg	Gln	Ala	Pro 25	Gly	Lys	Glu	Arg	Glu 30	Phe	Val
65	Ser	Xaa	Xaa 35	Arg	Phe	Thr	Ile	Ser 40	Gln	Asp	Ser	Ala	Lys 45	Asn	Thr	Val

	Tyr	Leu 6	Gln	Met	Asn	Ser	Leu 55	Lys	Pro	Glu	Asp	Thr 60	Ala	Met	Tyr.	Tyr
5	Cys 65	Lys	Ile	Xaa	Xaa	Xaa 70	Trp	Gly	Gln	Gly	Thr 75	Gln	Val	Thr	Val	Ser 80
	Ser															
10	(2) INFO	RMATI	ON F	OR S	SEQ 1	מ מו): 7:	:								
15	(i)	(B) (C)	LEN TYP STR	IGTH: PE: & RANDE	ARACI 37 mino EDNES	amir aci SS: s	no ad id singl	cids								
	(ii)	MOLE	CULE	TYF	E: E	prote	ein									
20	(vii)				came	el "h		y cha shor						ent		
25	(xi)	SEQU	ENCE	DES	CRIE	OIT	N: SI	EQ II	NO:	: 7:						
	Trp	Gly	Gln	Gly	Thr 5	Gln	Val	Thr	Val	Ser 10	Ser	Gly	Thr	Asn	Glu 15	Val
30	Cys	Lys	Cys	Pro 20	Lys	Cys	Pro	Ala	Pro 25	Glu	Leu	Pro	Gly	Gly 30	Pro	Ser
	Val	Phe	Val 35	Phe	Pro											
35	(2) INFO	RMATI	ON F	FOR S	SEQ I	D NO): 8:	•								
40	(i)	(B) (C)	LEN TYP STP	GTH: PE: & RANDI	ARACT 60 amino EDNES	amin ac: SS: s	no ad id sing:	cids								
45	(ii)	MOLE	CULE	TYI	PE: 1	prote	ein									
15	(vii)				came	el "l		y cha long						nt		
5 0	(xi)	SEQU	ENCE	E DES	SCRI	PTIO	N: S1	EQ II	ои о	: 8:						
	Trp 1	Gly	Gln	Gly	Thr 5	Gln	Val	Thr	Val	Ser 10	Ser	Glu	Pro	Lys	Ile 15	Pro
5 5	Gln	Pro	Gln	Pro 20	Lys	Pro	Gln	Pro	Gln 25	Pro	Gln	Pro	Gln	Pro 30	Lys	Pro
60	Gln	Pro	Lys 35	Pro	Glu	Pro	Glu	Cys 40	Thr	Cys	Pro	Lys	Cys 45	Pro	Ala	Pro
OU	Glu	Leu 50	Leu	Gly	Gly	Pro	Ser 55	Val	Phe	Ile	Phe	Pro 60				

	(2) INFO	RMATION FOR SEQ ID NO: 9:
5	(i)	SEQUENCE CHARACTERISTICS: (A) LENGTH: 67 amino acids (B) TYPE: amino acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear
10	(ii)	MOLECULE TYPE: protein
10	(vii)	<pre>IMMEDIATE SOURCE: (B) CLONE: human gamma-3 CH1 - hinge - CH2 fragment</pre>
15	(xi)	SEQUENCE DESCRIPTION: SEQ ID NO: 9:
10	Lys 1	Val Asp Lys Arg Val Glu Leu Lys Thr Pro Leu Gly Asp Thr Thr 5 10 15
20	His	Thr Cys Pro Arg Cys Pro Glu Pro Lys Cys Ser Asp Thr Pro Pro 20 25 30
	Pro	Cys Pro Arg Cys Pro Glu Pro Lys Ser Cys Asp Thr Pro Pro Pro 35 40 45
25	Сув	Pro Arg Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Pho 50 55 60
30	Leu 65	Phe Pro
	(2) INFO	RMATION FOR SEQ ID NO: 10:
35	(i)	SEQUENCE CHARACTERISTICS: (A) LENGTH: 35 amino acids (B) TYPE: amino acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear
40	(ii)	MOLECULE TYPE: protein
	(vii)	IMMEDIATE SOURCE: (B) CLONE: human gamma-1 CH1 - hinge - CH2 fragment
45	(xi)	SEQUENCE DESCRIPTION: SEQ ID NO: 10:
	Lys 1	Val Asp Lys Lys Ala Glu Pro Lys Ser Cys Asp Lys Thr His The 5 10 15
50	Сув	Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Pho 20 25 30
55	Leu	Phe Pro 35
	(2) INFO	RMATION FOR SEQ ID NO: 11:
60	(i)	SEQUENCE CHARACTERISTICS: (A) LENGTH: 31 amino acids (B) TYPE: amino acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear
65	/;; \	MOI FOUL F TYPE: protein

	(vii)	IMME (B)	CLC	TE SO DNE:	OURCE huma	E: an ga	amma-	·2 CF	11 -	hing	re -	сн2	fraç	ment		
5	(xi)	SEQU	JENCE	E DES	SCRIE	OIT	1: SE	Q II	NO:	11:						
J	Lys 1	Val	Lys	Val	Thr 5	Val	Glu	Arg	Lys	Cys 10	Cys	Val	Glu	Cys	Pro 15	Pro
10	Cys	Pro	Ala	Pro 20	Pro	Val	Ala	Gly	Pro 25	Ser	Val	Phe	Leu	Phe 30	Pro	
	(2) INFO	RMATI	ON F	FOR S	SEQ 1	D NO): 12	: :								
15	(i)	(B)	LEN TYP STP	NGTH: PE: & RANDE	: 32 amino EDNES	TERIS amir aci SS: s	no ac id singl	ids								
2 0	(ii)	MOLE	CULE	E TYP	E: p	prote	ein									
25	(vii)					E: an ga	amma-	-4 CF	il -	hinç	je -	CH2	fraç	gment	:	
	(xi)	SEQU	JENCE	E DES	CRII	OITS	1: SE	Q II	NO:	12:						
30	Lys 1	Val	Asp	Lys	Arg 5	Val	Glu	Ser	Lys	Tyr 10	Gly	Pro	Pro	Сув	Pro 15	Ser
	Cys	Pro	Ala	Pro 20	Glu	Phe	Leu	Gly	Gly 25	Pro	Ser	Val	Phe	Leu 30	Phe	Pro
35	(2) INFO	RMATI	ON E	FOR S	SEQ I	ID NO): 13	3:								
4 0	(i)	(B)	LEN TYP STE	NGTH: PE: & RANDI	: 12: amino EDNES	reris lami o aci ss: s linea	ino a id singl	cids	5							
	(ii)	MOLE	CUL	E TYI	PE: 1	prote	ein									
4 5	(vii)					E: se he	eavy	cha:	in V-	-regi	Lon					
		SEQU														
5 0	1	Val			5					10					15	
<i>5</i> 5	Ser	Leu	Arg	Leu 20	Ser	Cys	Ala	Thr	Ser 25	Gly	Phe	Thr	Phe	Ser 30	Asp	Phe
	Tyr	Met	Glu 35	Trp	Val	Arg	Gln	Pro 40	Pro	Gly	Lys	Arg	Leu 45	Glu	Trp	Ile
60	Ala	Ala 50	Ser	Arg	Asn	Lys	Ala 55	Asn	Asp	Tyr	Thr	Thr 60	Glu	Tyr	Ser	Ala
	Ser 65	Val	Lys	Gly	Arg	Phe 70	Ile	Val	Ser	Arg	Asp 75	Thr	Ser	Gln	Ser	Ile 80
65	Leu	Tyr	Leu	Gln	Met 85	Asn	Ala	Leu	Arg	Ala 90	Glu	Asp	Thr	Ala	Ile 95	Tyr

							•									
	Tyr	Cys	Ala	Ar g 10 0	Asp	Tyr	Tyr	Gly	Ser 105	Ser	Tyr	Phe	Asp	Val 110	Trp	Gly
5	Ala	Gly	Thr 115	Thr	Val	Thr	Val	Ser 120	Ser							
	(2) INFO	RMAT:	ION P	FOR S	SEQ :	ID NO): 1	4:								
10	(i)	(A (B	UENCI	NGTH PE: 8	: 13: amino	l am:	ino a id	acid	5							
15) STI) TOI					ıe								
	(ii)		•		_		≘in									
20	(vii)		EDIA:				eavy	cha	in V	-reg	ion					
	(xi)	SEQ	UENCI	E DES	SCRII	PTIO	N: S1	EQ II	D NO	: 14	:					
25	Glu 1	Val	Gln	Leu	Val 5	Glu	Ser	Gly	Gly	Gly 10	Leu	Val	Gln	Pro	Gly 15	Gly
	Ser	Leu	Arg	Leu 20	Ser	Cys	Ala	Ala	Ser 25	Gly	Phe	Thr	Phe	Ser 30	Ser	Tyr
30	Ala	Met	Ser 35	Trp	Val	Arg	Gln	Ala 40	Pro	Gly	Lys	Gly	Leu 45	Glu	Trp	Val
	Ser	Xaa 50	Ile	Ser	Xaa	Lys	Thr 55	Asp	Gly	Gly	Xaa	Thr 60	Tyr	Tyr	Ala	Asp
35	Ser 65	Val	Lys	Gly	Arg	Phe 70	Thr	Ile	Ser	Arg	Asp 75	Asn	Ser	Lys	Asn	Thr 80
40	Leu	Tyr	Leu	Gln	Met 85	Asn	Ser	Leu	Arg	Ala 90	Glu	Asp	Thr	Ala	Val 95	Tyr
70	Tyr	Cys	Ala	A rg 10 0	Xaa	Xaa	Xaa	Xaa	Xaa 105	Xaa	Xaa	Xaa	Xaa	Xaa 110	Xaa	Tyr
45	Tyr	Tyr	Tyr 115	His	Xaa	Phe	Asp	Tyr 120	Trp	Gly	Gln	Gly	Thr 125	Leu	Val	Thr
	Val	Ser 130	Ser													
50	(2) INFO	RMAT	ION 1	FOR :	SEQ :	ID NO	o: 1	5:								
55	(i)	(A (B (C	UENCI) LEI) TYI) STI) TOI	NGTH PE: A RAND	: 11 amino EDNE:	4 am: o ac: SS: :	ino a id sing	acid	s							
60	(ii)	MOL	ECULI	E TY	PE:]	prot	ein									
	(vii)		EDIA:				heav	y ch	ain .	im mu:	nogl	obul	in"	V-re	gion	(1)
65	(xi)	SEQ	UENCI	E DE	SCRI	PTIO	N: 5	EQ I	D NO	: 15	:					
~~	Gly 1	Gly	Ser	Val	Gln 5	Ala	Gly	Gly	Ser	Leu 10	Arg	Leu	Ser	Cys	Ala 15	Ala

Ser	Gly	Tyr	Ser 20	Asn	Cys	Pro	Leu	Thr 25	Trp	Ser	Trp	Tyr	Arg 30	Glņ	Phe
Pro	Gly	Thr 35	Glu	Arg	Glu	Phe	Val 40	Ser	Ser	Met	Asp	Pro 45	Asp	Gly	Asn
Thr	Lys 50	Tyr	Thr	Tyr	Ser	Val 55	Lys	Gly	Arg	Phe	Thr 60	Met	Ser	Arg	Gly
Ser 65	Thr	Glu	Tyr	Thr	Val 70	Phe	Leu	Gln	Met	Asp 75	Asn	Leu	Lys	Pro	Gl u 80
Asp	Thr	Ala	Met	T yr 85	Tyr	Cys	Lys	Thr	Ala 90	Leu	Gln	Pro	Gly	Gly 95	Tyr
Cys	Gly	Tyr	Gly 100	Xaa	Cys	Leu	Trp	Gly 105	Gln	Gly	Thr	Gln	Val 110	Thr	Val
Ser	Ser														
(2) INFO	RMAT:	I NOI	FOR S	SEQ I	ID NO): 16	5:								
(i)	(A (B (C) LEI) TYI) STI	NGTH: PE: & RANDI	: 120 amino EDNES	D ami D aci	ino a id sing!	cids	5							
(ii)	MOL	ECULI	E TY	PE: p	prote	ein									
(vii)						neavy	y cha	ain i	immu	oglo	bul:	in" \	/-re	gion	(2)
(xi)	SEQ	UENCI	E DE	SCRII	PTIO	1: SI	EQ II	ON C	: 16:	:					
Asp 1	Val	Gln	Leu	Val 5	Ala	Ser	Gly	Gly	Gly 10	Ser	Val	Gln	Ala	Gly 15	Gly
Ser	Leu	Arg	Leu 20	Ser	Cys	Thr	Ala	Ser 25	Gly	Asp	Ser	Phe	ser 30	Arg	Phe
Ala	Met	Ser 35	Trp	Phe	Arg	Gln	Ala 40	Pro	Gly	Lys	Glu	Cys 45	Glu	Leu	Val
Ser	Ser 50	Ile	Gln	Ser	Asn	Gly 55	Arg	Thr	Thr			Asp	Ser	Val	Gln
Gly 65	Arg	Phe	Thr	Ile	ser 70	Arg	Asp	Asn	Ser	Arg 75	Asn	Thr	Val	Tyr	Leu 80
Gln	Met	Asn	Ser	Leu 85	Lys	Pro	Glu	Asp	Thr 90	Ala	Val	Tyr	Tyr	Cys 95	Gly
Ala	Val	Ser	Leu 100	Met	Asp	Arg	Ile	Ser 105	Gln	His	Gly	Cys	Arg 110	Gly	Gln
Gly	Thr	Gln 115	Val	Thr	Val	Ser	Leu 120								
12) THEO															
(2) INFO	RMAT	ION :	FOR :	SEQ	ID NO	J: 1	7:								
	Pro Thr Ser 65 Asp Cys Ser (2) INFO (ii) (vii) (xi) Asp 1 Ser Ala Ser Gly 65 Gln Ala	Pro Gly Thr Lys 50 Ser Thr 65 Asp Thr Cys Gly Ser Ser (2) INFORMAT: (i) SEQN (A (B (C) (D (ii) MOLN (vii) IMMN (B (xi) SEQN Asp Val 1 Ser Leu Ala Met Ser Ser 50 Gly Arg 65 Gln Met Ala Val	Pro Gly Thr 35 Thr Lys Tyr 50 Ser Thr Glu 65 Asp Thr Ala Cys Gly Tyr Ser Ser (2) INFORMATION 1 (i) SEQUENCY (A) LEY (B) TYY (C) STY (D) TOY (B) CLC (C) STY (D) TOY (B) CLC (C) SEQUENCY (C) SEQUE	Pro Gly Thr Glu 35 Thr Lys Tyr Thr 50 Ser Thr Glu Tyr 65 Asp Thr Ala Met Cys Gly Tyr Gly 100 Ser Ser (2) INFORMATION FOR S (1) SEQUENCE CHA (A) LENGTH: (B) TYPE: 2 (C) STRANDE (D) TOPOLOG (ii) MOLECULE TYR (vii) IMMEDIATE SC (B) CLONE: (xi) SEQUENCE DES Asp Val Gln Leu 1 Ser Leu Arg Leu 20 Ala Met Ser Trp 35 Ser Ser Ile Gln 50 Gly Arg Phe Thr 65 Gln Met Asn Ser Ala Val Ser Leu 100 Gly Thr Gln Val	Pro Gly Thr Glu Arg 35 Thr Lys Tyr Thr Tyr 50 Ser Thr Glu Tyr Thr 65 Asp Thr Ala Met Tyr 85 Cys Gly Tyr Gly Xaa 100 Ser Ser (2) INFORMATION FOR SEQ (i) SEQUENCE CHARACT (A) LENGTH: 12 (B) TYPE: amino (C) STRANDEDNES (D) TOPOLOGY: (ii) MOLECULE TYPE: p (vii) IMMEDIATE SOURCE (B) CLONE: came (xi) SEQUENCE DESCRIM Asp Val Gln Leu Val 1 Ser Leu Arg Leu Ser 20 Ala Met Ser Trp Phe 35 Ser Ser Ile Gln Ser 50 Gly Arg Phe Thr Ile 65 Gln Met Asn Ser Leu 85 Ala Val Ser Leu Met 100 Gly Thr Gln Val Thr	Pro Gly Thr Glu Arg Glu 35 Thr Lys Tyr Thr Tyr Ser 50 Ser Thr Glu Tyr Thr Val 65 Asp Thr Ala Met Tyr Tyr 85 Cys Gly Tyr Gly Xaa Cys 100 Ser Ser (2) INFORMATION FOR SEQ ID NO (i) SEQUENCE CHARACTERIS (A) LENGTH: 120 ami (B) TYPE: amino aci (C) STRANDEDNESS: s (D) TOPOLOGY: lines (ii) MOLECULE TYPE: prote (vii) IMMEDIATE SOURCE: (B) CLONE: camel "No (xi) SEQUENCE DESCRIPTION Asp Val Gln Leu Val Ala 1 Ser Leu Arg Leu Ser Cys 20 Ala Met Ser Trp Phe Arg 35 Ser Ser Ile Gln Ser Asn 50 Gly Arg Phe Thr Ile Ser 65 Gly Arg Phe Thr Ile Ser 65 Ala Val Ser Leu Met Asp 100 Gly Thr Gln Val Thr Val	Pro Gly Thr Glu Arg Glu Phe 35 Thr Lys Tyr Thr Tyr Ser Val 50 Ser Thr Glu Tyr Thr Val Phe 65 Asp Thr Ala Met Tyr Tyr Cys 85 Cys Gly Tyr Gly Xaa Cys Leu 100 Ser Ser (2) INFORMATION FOR SEQ ID NO: 16 (i) SEQUENCE CHARACTERISTICS (A) LENGTH: 120 amino acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear (ii) MOLECULE TYPE: protein (vii) IMMEDIATE SOURCE: (B) CLONE: camel "heavy (xi) SEQUENCE DESCRIPTION: SI Asp Val Gln Leu Val Ala Ser 1 5 Ser Leu Arg Leu Ser Cys Thr 20 Ala Met Ser Trp Phe Arg Gln 35 Ser Ser Ile Gln Ser Asn Gly 50 Gly Arg Phe Thr Ile Ser Arg 65 Gly Arg Phe Thr Ile Ser Arg 65 Ala Val Ser Leu Met Asp Arg 100 Gly Thr Gln Val Thr Val Ser	Pro Gly Thr Glu Arg Glu Phe Val 35	Pro Gly Thr Glu Arg Glu Phe Val Ser 35 Thr Lys Tyr Thr Tyr Ser Val Lys Gly 55 Ser Thr Glu Tyr Thr Val Phe Leu Gln 70 Asp Thr Ala Met Tyr Tyr Cys Lys Thr 85 Cys Gly Tyr Gly Xaa Cys Leu Trp Gly 100 Ser Ser (2) INFORMATION FOR SEQ ID NO: 16: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 120 amino acids (B) TYPE: amino acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear (ii) MOLECULE TYPE: protein (vii) IMMEDIATE SOURCE: (B) CLONE: camel "heavy chain is sequence Description: SEQ ID No: Asp Val Gln Leu Val Ala Ser Gly Gly 1 Ser Leu Arg Leu Ser Cys Thr Ala Ser 20 Ala Met Ser Trp Phe Arg Gln Ala Pro 35 Gly Arg Phe Thr Ile Ser Asn Gly Arg Thr 50 Gly Arg Phe Thr Ile Ser Arg Asp Asn 65 Ala Val Ser Leu Met Asp Arg Ile Ser 100 Gly Thr Gln Val Thr Val Ser Leu	Pro Gly Thr Glu Arg Glu Phe Val Ser Ser 35 Thr Lys Tyr Thr Tyr Ser Val Lys Gly Arg 50 Ser Thr Glu Tyr Thr Val Phe Leu Gln Met 70 Asp Thr Ala Met Tyr Tyr Cys Lys Thr Ala 85 Cys Gly Tyr Gly Xaa Cys Leu Trp Gly Gln 100 Ser Ser (2) INFORMATION FOR SEQ ID NO: 16: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 120 amino acids (B) Type: amino acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear (ii) MOLECULE TYPE: protein (vii) IMMEDIATE SOURCE: (B) CLONE: camel "heavy chain immur (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16: Asp Val Gln Leu Val Ala Ser Gly Gly Gly 10 Ser Leu Arg Leu Ser Cys Thr Ala Ser Gly 20 Ala Met Ser Trp Phe Arg Gln Ala Pro Gly 35 Gly Arg Phe Thr Ile Ser Arg Asp Asp Ser 70 Gln Met Asn Ser Leu Lys Pro Glu Asp Thr 85 Ala Val Ser Leu Met Asp Arg Ile Ser Gln 100 Gly Thr Gln Val Thr Val Ser Leu	Pro Gly Thr Glu Arg Glu Phe Val Ser Ser Met 35 Thr Lys Tyr Thr Tyr Ser Val Lys Gly Arg Phe 50 Ser Thr Glu Tyr Thr Val Phe Leu Gln Met Asp 75 Asp Thr Ala Met Tyr Tyr Cys Lys Thr Ala Leu 85 Cys Gly Tyr Gly Xaa Cys Leu Trp Gly Gln Gly 100 Ser Ser (2) INFORMATION FOR SEQ ID NO: 16: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 120 amino acids (B) TYPE: amino acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear (ii) MOLECULE TYPE: protein (vii) IMMEDIATE SOURCE: (B) CLONE: camel "heavy chain immunogld (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16: Asp Val Gln Leu Val Ala Ser Gly Gly Gly Ser 10 Ser Leu Arg Leu Ser Cys Thr Ala Ser Gly Asp 20 Ala Met Ser Trp Phe Arg Gln Ala Pro Gly Lys 35 Gly Arg Phe Thr Ile Ser Arg Asp Asn Ser Arg 70 Gln Met Asn Ser Leu Lys Pro Glu Asp Thr Ala 85 Gly Arg Phe Thr Ile Ser Arg Asp Asn Ser Arg 70 Gln Met Asn Ser Leu Lys Pro Glu Asp Thr Ala 85 Gly Thr Gln Val Thr Val Ser Leu Gly Thr Gln His 100 Gly Thr Gln Val Thr Val Ser Leu	Pro Gly Thr Glu Arg Glu Phe Val Ser Ser Met Asp 35 Thr Lys Tyr Thr Tyr Ser Val Lys Gly Arg Phe Thr 50 Ser Thr Glu Tyr Thr Val Phe Leu Gln Met Asp Asn 65 Asp Thr Ala Met Tyr Tyr Cys Lys Thr Ala Leu Gln 85 Cys Gly Tyr Gly Xaa Cys Leu Trp Gly Gln Gly Thr 100 Ser Ser (2) INFORMATION FOR SEQ ID NO: 16: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 120 amino acids (B) TYPE: amino acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear (ii) MOLECULE TYPE: protein (vii) IMMEDIATE SOURCE: (B) CLONE: camel "heavy chain immunoglobul: (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16: Asp Val Gln Leu Val Ala Ser Gly Gly Gly Ser Val 1 Ser Leu Arg Leu Ser Cys Thr Ala Ser Gly Asp Ser 20 Ala Met Ser Trp Phe Arg Gln Ala Pro Gly Lys Glu 35 Gly Arg Phe Thr Ile Ser Arg Asp Asn Ser Arg Asn 65 Gly Arg Phe Thr Ile Ser Arg Asp Asn Ser Arg Asn 65 Gln Met Asn Ser Leu Lys Pro Glu Asp Thr Ala Val 85 Ala Val Ser Leu Met Asp Arg Ile Ser Gln His Gly 100 Gly Thr Gln Val Thr Val Ser Leu	Pro Gly Thr Glu Arg Glu Phe Val Ser Ser Met Asp Pro 40 Ser Thr Lys Tyr Thr Tyr Ser Val Lys Gly Arg Phe Thr Met 50 Ser Thr Glu Tyr Thr Val Phe Leu Gln Met Asp Asn Leu 70 Phe Asp Thr Ala Met Tyr Tyr Cys Lys Thr Ala Leu Gln Pro 85 Ser Ser Ser Ser Ser Ser Thr Gly Xaa Cys Leu Trp Gly Gln Gly Thr Gln 100 Ser Ser (a) Length: 120 amino acids (b) TypE: amino acid (c) STRANDEDNESS: single (d) Topolocy: linear (ii) MOLECULE TypE: protein (vii) IMMEDIATE SOURCE: (B) CLONE: camel "heavy chain immunoglobulin" (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16: Asp Val Gln Leu Val Ala Ser Gly Gly Gly Ser Val Gln 1 Ser Leu Arg Leu Ser Cys Thr Ala Ser Gly Asp Ser Phe 20 Ala Met Ser Trp Phe Arg Gln Ala Pro Gly Lys Glu Cys 35 Ser Ser Ile Gln Ser Asn Gly Arg Thr Thr Glu Ala Asp 50 Gly Arg Phe Thr Ile Ser Arg Asp Asn Ser Arg Asn Thr 70 Gln Met Asn Ser Leu Lys Pro Glu Asp Thr Ala Val Tyr 85 Gln Met Asn Ser Leu Lys Pro Glu Asp Thr Ala Val Tyr 85 Gly Thr Gln His Gly Cys 100 Gly Thr Gln Val Thr Val Ser Leu	Pro Gly Thr Glu Arg Glu Phe Val Ser Ser Met Asp Pro Asp 45 Thr Lys Tyr Thr Tyr Ser Val Lys Gly Arg Phe Thr Met Ser 50 Ser Thr Glu Tyr Thr Val Phe Leu Gln Met Asp Asn Leu Lys 75 Asp Thr Ala Met Tyr Tyr Cys Lys Thr Ala Leu Gln Pro Gly 85 Cys Gly Tyr Gly Xaa Cys Leu Trp Gly Gln Gly Thr Gln Val 100 Ser Ser (2) INFORMATION FOR SEQ ID NO: 16: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 120 amino acids (B) TYPE: amino acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear (ii) MOLECULE TYPE: protein (vii) IMMEDIATE SOURCE: (B) CLONE: camel "heavy chain immunoglobulin" V-reg (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16: Asp Val Gln Leu Val Ala Ser Gly Gly Gly Ser Val Gln Ala 1 Ser Leu Arg Leu Ser Cys Thr Ala Ser Gly Asp Ser Phe Ser 20 Ala Met Ser Trp Phe Arg Gln Ala Pro Gly Lys Glu Cys Glu 35 Ser Ser Ile Gln Ser Asn Cly Arg Thr Thr Glu Ala Asp Ser 50 Gly Arg Phe Thr Ile Ser Arg Asp Asn Ser Arg Asn Thr Val 65 Gln Met Asn Ser Leu Lys Pro Glu Asp Thr Ala Val Tyr Tyr 85 Ala Val Ser Leu Met Asp Arg Ile Ser Gln His Gly Cys Arg 100 Gly Thr Gln Val Thr Val Ser Leu	Pro Gly Thr Glu Arg Glu Phe Val Ser Ser Met Asp Pro Asp Gly Thr Lys Tyr Thr Tyr Ser Val Lys Gly Arg Phe Thr Met Ser Arg Ser Thr Glu Tyr Thr Val Phe Leu Gln Met Asp Asn Leu Lys Pro 65 Asp Thr Ala Met Tyr Tyr Cys Lys Thr Ala Leu Gln Pro Gly Gly 85 Cys Gly Tyr Gly Xaa Cys Leu Trp Gly Gln Gly Thr Gln Val Thr 100 Ser Ser (2) INFORMATION FOR SEQ ID NO: 16: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 120 amino acids (C) STRANDEDNESS: single (D) TOPOLOGY: linear (ii) MOLECULE TYPE: protein (vii) IMMEDIATE SOURCE: (B) CLONE: camel "heavy chain immunoglobulin" V-region (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16: Asp Val Gln Leu Val Ala Ser Gly Gly Gly Ser Val Gln Ala Gly 1 Ser Leu Arg Leu Ser Cys Thr Ala Ser Gly Asp Ser Phe Ser Arg 20 Ala Met Ser Trp Phe Arg Gln Ala Pro Gly Lys Glu Cys Glu Leu 35 Ser Ser Ile Gln Ser Asn Gly Arg Thr Thr Glu Ala Asp Ser Val 50 Gly Arg Phe Thr Ile Ser Arg Asp Asn Ser Arg Asn Thr Val Tyr 65 Gln Met Asn Ser Leu Lys Pro Glu Asp Thr Ala Val Tyr Tyr Cys 85 Ala Val Ser Leu Met Asp Arg Ile Ser Gln His Gly Cys Arg Gly Gly Thr Gln Val Thr Val Ser Leu

		(D) TOI	POLO	3Y: 3	linea	ar									
	(ii)	MOLI	ECULI	E TYI	PE: I	prote	ein						•			
5	(vii)						neavy	/ cha	ain i	immur	noglo	bul:	in" V	/-rec	gion	(3)
	(xi)	SEQ	JENCI	E DES	SCRI	OITS	N: SE	EQ II	NO:	: 17:	:					
10	Gly 1	Gly	Ser	Val	Gln 5	Thr	Gly	Gly	Ser	Leu 10	Arg	Leu	Ser	Cys	Ala 15	۷a
15	Ser	Gly	Phe	S er 20	Phe	Ser	Thr	Ser	Cys 25	Met	Ala	Trp	Phe	Arg 30	Gln	Ala
10	Ser	Gly	Lys 35	Gln	Arg	Glu	Gly	Val 40	Ala	Ala	Ile	Asn	Ser 45	Gly	Gly	Gly
2 0	Arg	Thr 50	Tyr	Tyr	Asn	Thr	Tyr 55	Val	Ala	Glu	Ser	Val 60	Lys	Gly	Arg	Phe
	Ala 65	Ile	Ser	Gln	Asp	Asn 70	Ala	Lys	Thr	Thr	Val 75	Tyr	Leu	Asp	Met	Ası 80
25	Asn	Leu	Thr	Pro	Glu 85	Asp	Thr	Ala	Thr	Tyr 90	Tyr	Сув	Ala	Ala	Val 95	Pro
30	Ala	His	Leu	Gly 100	Pro	Gly	Ala	Ile	Leu 105	Asp	Leu	Lys	Lys	Tyr 110	Lys	Туз
50	Trp	Gly	Gln 115	Gly	Thr	Gln	Val	Thr 120	Val	Ser	Ser					
35	(2) INFO	RMAT:	ION I	FOR S	SEQ :	D NO	D: 18	3:								
4 0	(i)	(B) LEI) TYI	NGTH	: 116 amino	am:	ino a id		3							
		(D)	TOI	POLO	3Y: :			le								
	(ii)	•) TOI			linea	ar	le								
45	(ii) (vii)	MOLI) TOI ECULI EDIA:	E TYI	PE: p	linea prote E:	ar ein		ain :	immur	noglo	obul:	in" V	/-rec	gion	(7)
45		MOLI	TOIECULI	E TYI FE SO ONE:	PE: p OURCI	linea prote E: el "l	ar ein neavy	y cha			-	obul:	in" V	/-rec	jion	(7)
45 50	(vii)	MOLI IMMI (B SEQI	ECULI EDIA: CLO UENCI	E TYI PE SO ONE: E DE:	PE: p OURCE Came	linea prote E: El "l	ar ein neavy N: SH	/ Cha	ONO:	: 18:	:			-	gion Ala 15	•
50	(vii) (xi) Gly 1	MOLI IMMI (B SEQI Gly	ECULI ECULI EDIA:) CLC UENCI	E TYI TE SC ONE: E DES	PE: p DURCE came SCRII Gln 5	linea prote 3: el "l PTION Gly	ein neavy N: SH Gly	y cha EQ II Gly	NO:	: 18: Leu 10	: Arg	Leu	Ser	Cys	Ala	Ile
	(vii) (xi) Gly 1 Ser	MOLI IMMI (B SEQI Gly	ECULI ECULI EDIA:) CLU UENCI Ser Tyr	TYIONE: Val Thr	PE: POURCE came SCRIP Gln 5	linea prote 3: el "l PTION Gly	ein neavy N: SH Gly Ser	y cha EQ II Gly Phe	Ser Cys 25	Leu 10 Met	Arg Gly	Leu Trp	Ser Phe	Cys Arg 30	Ala 15	Ile
50	(vii) (xi) Gly 1 Ser	MOLI IMMI (B SEQI Gly Gly	ECULION CLUENCI Ser Tyr Lys 35	TE SCONE: E DES Val Thr 20 Glu	PE: I DURCE came SCRII 5 Gln 5 Tyr	linea prote E: el "l PTION Gly Gly	ein neavy N: SH Gly Ser Gly	cha Gly Phe Ile 40	Ser Cys 25	Leu 10 Met	Arg Gly	Leu Trp Leu	Ser Phe Asn 45	Cys Arg 30 Gly	Ala 15 Glu	Ile Gly
50 55	(vii) (xi) Gly 1 Ser Pro	MOLI IMMI (B SEQI Gly Gly Gly	ECULIANDE CLUENCIANDE SET Tyr Lys 35 Tyr	TE SOONE: E DES Val Thr 20 Glu	PE: IDURCE CAME SCRII Gln 5 Tyr Arg	PTION Gly Gly Asp	ein neavy N: SI Gly Ser Gly Ser	Cha CQ II Gly Phe Ile 40 Val	Ser Cys 25 Ala	Leu 10 Met Thr	Arg Gly Ile	Leu Trp Leu Phe 60	Ser Phe Asn 45 Thr	Cys Arg 30 Gly	Ala 15 Glu Gly	Ile Gly The

	Cys	Glu	Leu	Pro 100	Leu	Leu	Phe	Asp	Tyr 105	Trp	Gly	Gln	Gly	Thr 110	Gln,	Val
5	Thr	Val	Ser 115	Ser												
	(2) INFO	RMAT:	ION I	FOR S	SEQ :	ID NO	o: 19):								
10	(i)	(A (B (C) LEI) TYI) STI	E CHANGTH: PE: & RANDI	: 114 amino EDNES	ami aci	ino a id sing:	acid	5							
15	(ii)		•	POLOC TYI			_									
•	(vii)	IMMI	EDIA:		DURCI	:		y cha	ain :	immur	noglo	bul:	in" V	/-red	gion	(9)
20	(xi)	SEQ	JENCI	E DES	SCRII	OITS	1: SI	EQ II	ON C	: 19:	:					
25	Gly 1	Gly	Ser	Val	Gln 5	Ala	Gly	Gly	Ser	Leu 10	Thr	Leu	Ser	Cys	Val 15	Tyr
43	Thr	Asn	yab	Thr 20	Gly	Thr	Met	Gly	Trp 25	Phe	Arg	Gln	Ala	Pro 30	Gly	Lys
30	Glu	Cys	Glu 35	Arg	Val	Ala	His	Ile 40	Thr	Pro	Asp	Gly	Met 45	Thr	Phe	Ile
	Asp	Glu 50	Pro	Val	Lys	Gly	Arg 55	Phe	Thr	Ile	Ser	Arg 60	Asp	Asn	Ala	Gln
35	Lys 65	Thr	Leu	Ser	Leu	Arg 70	Met	Asn	Ser	Leu	Arg 75	Pro	Glu	Asp	Thr	Ala 80
	Val	Tyr	Tyr	Cys	Ala 85	Ala	Asp	Trp	Lys	Tyr 90	Trp	Thr	Cys	Gly	Ala 95	Gln
40	Thr	Gly	Gly	Tyr 100	Phe	Gly	Gln	Trp	Gly 105	Gln	Gly	Ala	Gln	Val 110	Thr	Val
45	Ser	Ser														
	(2) INFO	RMAT:	ION I	FOR :	SEQ :	ID NO	D: 20	0:								
50	(i)	(A (B (C) LEI) TYI) STI	E CHI NGTH PE: 7 RANDI POLO	: 12 amino EDNE	5 am: 5 ac: 5S: 1	ino a id sing:	acid	5							
55	(ii)	MOL	ECULI	E TY	PE:]	prote	ein									
	(vii)			TE SO			heav	y ch	ain .	immu	nogl	obul	in" '	V-re	gion	(11
60	(xi)	SEQ	UENC	E DE	SCRI	PTIO	N: S	EQ I	on o	: 20	:					
	Gly 1	Gly	Ser	Val	Gln 5	Ala	Gly	Gly	Ser	Leu 10	Arg	Leu	Ser	Cys	Asn 15	Val
65	Ser	Gly	Ser	Pro 20	Ser	Ser	Thr	Tyr	Cys 25	Leu	Gly	Trp	Phe	Arg 30	Gln	Ala

	Pro	Gly	Arg 35	Glu	Arg	Glu	Gly	Val 40	Thr	Ala	Ile	Asn	Thr 45	Asp	Gly.	Ser
5	Ile	Ile 50	Tyr	Ala	Ala	Asp	Ser 55	Val	Lys	Gly	Arg	Phe 60	Thr	Ile	Ser	Gln
	Asp 65	Thr	Ala	Lys	Glu	Thr 70	Val	His	Leu	Gln	Met 75	Asn	Asn	Leu	Gln	Pro 80
10	Glu	Asp	Thr	Ala	Thr 85	Tyr	Tyr	Cys	Ala	Ala 90	Arg	Leu	Thr	Glu	Met 95	Gly
15	Ala	Cys	Asp	Ala 100	Arg	Trp	Ala	Thr	Leu 105	Ala	Thr	Arg	Thr	Phe 110	Ala	Tyr
	Asn	Tyr	Trp 115	Gly	Gln	Gly	Thr	Gln 120	Val	Thr	Val	Ser	Ser 125			
20	(2) INFO	RMAT	ION I	FOR S	SEQ :	ID NO	D: 2	1:								
25	(i)	(A (B (C	LEI TYI STI	IGTH: PE: 3 RANDI	: 114 amino EDNES	TERIS am: ac: SS: s	ino a id sing:	acids	5							
	(ii)	MOLI	ECULI	TYI	PE: 1	prote	∍in				•					
30	(vii)						neavy	y cha	ain :	immu	noglo	bul:	in" \	/-re	gion	(13
	(xi)	SEQ	JENCI	DES	SCRII	PTIO	N: SI	EQ II	ON C	: 21:	:					
35				Val	Glu 5	Ala	Gly	Gly	Ser	Leu 10	Arg	Leu	Ser	Сув	Thr 15	Ala
35 40	Gly 1	Gly	Ser		5			_	Ser Trp 25	10	_				15	
	Gly 1 Ser	Gly Gly Arg	Ser Tyr Glu 35	Val 20 Gly	5 Ser Val	Ser Ala	Met Phe	Ala Val 40	Trp 25 Gln	10 Phe Thr	Arg Ala	Gln Asp	Val Asn 45	Pro 30 Ser	15 Gly Ala	Gln Leu
	Gly 1 ser Glu Tyr	Gly Gly Arg Gly 50	Ser Tyr Glu 35 Asp	Val 20 Gly Ser	5 Ser Val	Ser Ala Lys	Met Phe Gly 55	Ala Val 40 Arg	Trp 25 Gln Phe	10 Phe Thr	Arg Ala Ile	Gln Asp Ser 60	Val Asn 45 His	Pro 30 Ser	15 Gly Ala Asn	Gln Leu Ala
40	Gly 1 ser Glu Tyr	Gly Gly Arg Gly 50	Ser Tyr Glu 35 Asp	Val 20 Gly Ser	5 Ser Val	Ser Ala Lys	Met Phe Gly 55	Ala Val 40 Arg	Trp 25 Gln	10 Phe Thr	Arg Ala Ile	Gln Asp Ser 60	Val Asn 45 His	Pro 30 Ser	15 Gly Ala Asn	Gln Leu Ala
40	Gly 1 ser Glu Tyr Lys 65	Gly Gly Arg Gly 50 Asn	Ser Tyr Glu 35 Asp	Val 20 Gly Ser Leu	5 Ser Val Val	Ser Ala Lys Leu 70	Met Phe Gly 55 Gln	Ala Val 40 Arg	Trp 25 Gln Phe	10 Phe Thr Thr	Arg Ala Ile Leu 75	Gln Asp Ser 60 Gln	Val Asn 45 His	Pro 30 Ser Asp	15 Gly Ala Asn Asp	Gln Leu Ala Thr
40 45 50	Gly 1 Ser Glu Tyr Lys 65 Gly	Gly Gly Arg Gly 50 Asn Val	Ser Tyr Glu 35 Asp Thr	Val 20 Gly Ser Leu	Ser Val Val Tyr Cys 85	Ser Ala Lys Leu 70 Ala	Met Phe Gly 55 Gln Ala	Ala Val 40 Arg Met	Trp 25 Gln Phe Arg	10 Phe Thr Thr Asn Lys 90	Arg Ala Ile Leu 75 Asp	Gln Asp Ser 60 Gln Arg	Val Asn 45 His Pro	Pro 30 Ser Asp Asp	Gly Ala Asn Asp Trp 95	Gln Leu Ala Thr 80 Ala
40 45	Gly 1 Ser Glu Tyr Lys 65 Gly	Gly Gly Arg Gly 50 Asn Val	Ser Tyr Glu 35 Asp Thr	Val 20 Gly Ser Leu Tyr	Ser Val Val Tyr Cys 85	Ser Ala Lys Leu 70 Ala	Met Phe Gly 55 Gln Ala	Ala Val 40 Arg Met	Trp 25 Gln Phe Arg Lys	10 Phe Thr Thr Asn Lys 90	Arg Ala Ile Leu 75 Asp	Gln Asp Ser 60 Gln Arg	Val Asn 45 His Pro	Pro 30 Ser Asp Asp	Gly Ala Asn Asp Trp 95	Gln Leu Ala Thr 80 Ala
40 45 50 55	Gly 1 Ser Glu Tyr Lys 65 Gly	Gly Gly Arg Gly 50 Asn Val Pro	Ser Tyr Glu 35 Asp Thr Tyr	Val 20 Gly Ser Leu Tyr Glu 100	Ser Val Val Tyr Cys 85 Trp	Ser Ala Lys Leu 70 Ala Asn	Met Phe Gly 55 Gln Ala Asn	Ala Val 40 Arg Met Gln Trp	Trp 25 Gln Phe Arg Lys	10 Phe Thr Thr Asn Lys 90	Arg Ala Ile Leu 75 Asp	Gln Asp Ser 60 Gln Arg	Val Asn 45 His Pro	Pro 30 Ser Asp Asp	Gly Ala Asn Asp Trp 95	Gln Leu Ala Thr 80 Ala
40 45 50	Gly 1 Ser Glu Tyr Lys 65 Gly Glu Ser (2) INFO	Gly Gly Arg Gly 50 Asn Val Pro Ser RMAT: SEQUE (A	Ser Tyr Glu 35 Asp Thr Tyr Arg	Val 20 Gly Ser Leu Tyr Glu 100	Ser Val Val Tyr Cys 85 Trp SEQ ARAC:	Ser Ala Lys Leu 70 Ala Asn	Met Phe Gly 55 Gln Ala Asn O: 2: STIC: ino aid	Ala Val 40 Arg Met Gln Trp	Trp 25 Gln Phe Arg Lys Gly 105	10 Phe Thr Thr Asn Lys 90	Arg Ala Ile Leu 75 Asp	Gln Asp Ser 60 Gln Arg	Val Asn 45 His Pro	Pro 30 Ser Asp Asp	Gly Ala Asn Asp Trp 95	Gln Leu Ala Thr 80 Ala

(ii) MOLECULE TYPE: protein

	(vii)						neavy	cha	in i	immur	oglo	buli	ln" V	-rec	gion.	(16)
_	(xi)	SEQU	JENCE	E DES	CRIE	OIT	: SE	Q II	NO:	22:						
5	Gly 1	Gly	Ser	Ala	Gln 5	Ala	Gly	Gly	Ser	Leu 10	Arg	Leu	Ser	Cys	Ala 15	Ala
10	His	Gly	Ile	Pro 20	Leu	Asn	Gly	Tyr	Tyr 25	Ile	Ala	Trp	Phe	Arg 30	Gln	Ala
	Pro	Gly	Lys 35	Gly	Arg	Glu	Gly	Val 40	Ala	Thr	Ile	Asn	Gly 45	Gly	Arg	Asp
15	Val	Thr 50	Tyr	Tyr	Ala	Asp	Ser 55	Val	Thr	Gly	Arg	Phe 60	Thr	Ile	Ser	Arg
2 0	Asp 65	Ser	Pro	Lys	Asn	Thr 70	Val	Tyr	Leu	Gln	Met 75	Asn	Ser	Leu	Lys	Pro 80
_0	Glu	Asp	Thr	Ala	Ile 85	Tyr	Phe	Cys	Ala	Ala 90	Gly	Ser	Arg	Phe	Ser 95	Ser
2 5	Pro	Val	Gly	Ser 100	Thr	Ser	Arg	Leu	Glu 105	Ser	Ser	Asp	Tyr	Asn 110	Tyr	Trp
	Gly	Gln	Gly 115	Ile	Gln	Val	Thr	Ala 120	Ser	Ser						
30	(2) INFO	RMATI	ION 1	FOR S	SEQ I	ED NO	o: 23	3:								
3 5	(i)	(A) (B) (C)	LEI TYI STI	E CHA NGTH: PE: & RANDI POLO	: 117 amino EDNES	7 am: 5 ac: 55: 5	ino a id singl	cida	3							
40	(ii)	MOLI	ECULI	E TYI	PE: I	prote	⊇in									
70	(Vii)						neavy	, cha	ain :	immur	noglo	bul:	in" (V-reç	gion	(17)
45	(xi)	SEQ	UENC	E DES	CRI	PTIO	V: SI	EQ II	ON C	23:	:					
10	Gly 1	Gly	Ser	Val	Gln 5	Pro	Gly	Gly	Ser	Leu 10	Thr	Leu	Ser	Cys	Thr 15	Val
50	Ser	Gly	Ala	Thr 20	Tyr	Ser	Asp	Tyr	Ser 25	Ile	Gly	Trp	Ile	Arg 30	Gln	Ala
	Pro	Gly	Lys 35	Asp	Arg	Glu	Val	Val 40	Ala	Ala	Ala	Asn	Thr 45	Gly	Ala	Thr
55	Ser	Lys 50	Phe	Tyr	Val	Asp	Phe 55	Val	Lys	Gly	Arg	Phe 60	Thr	Ile	Ser	Gln
60	Asp 65	Asn	Ala	Lys	Asn	Thr 70	Val	Tyr	Leu	Gln	Met 75	Ser	Phe	Leu	Lys	Pro 80
.,0	Glu	Asp	Thr	Ala	Ile 85	Tyr	Tyr	Сув	Ala	Ala 90	Ala	Asp	Pro	Ser	Ile 95	Tyr
65	Tyr	Ser	Ile	Leu	Xaa	Ile	Glu	Tyr	Lys	Tyr	Trp	Gly	Gln	Gly		Gln

Val Thr Val Ser Ser 115

5	(2) INFO	RMATI	ON F	OR S	SEQ 1	ID NO	D: 24	1:								
10	(i)	(B) (C)	LEN TYP STR	IGTH: PE: & RANDE	: 123 amino EDNES	am:	ino a id sing!	acid	3							
	(ii)	MOLE	CULE	TYF	E: F	orote	ein									
15	(vii)						neavy	, cha	ain :	immur	noglo	bul i	in" V	/-rec	gion	(18)
	(xi)	SEQU	ENCE	DES	SCRIE	OIT	N: SI	EQ II	ON C	: 24:	:					
20	Gly 1	Gly	Ser	Val	Gln 5	Ala	Gly	Gly	Ser	Leu 10	Arg	Leu	Ser	Cys	Thr 15	Gly
25	Ser	Gly	Phe	Pro 20	Tyr	Ser	Thr	Phe	Cys 25	Leu	Gly	Trp	Phe	Arg 30	Gln	Ala
25	Pro	Gly	Lys 35	Glu	Arg	Glu	Gly	Val 40	Ala	Gly	Ile	Asn	Ser 45	Ala	Gly	Gly
30	Asn	Thr 50	Tyr	Tyr	Ala	Asp	Ala 55	Val	Lys	Gly	Arg	Phe 60	Thr	Ile	Ser	Gln
	Gly 65	Asn	Ala	Lys	Asn	Thr 70	Val	Phe	Leu	Gln	Met 75	Asp	Asn	Leu	Lys	Pro 80
35	Glu	Asp	Thr	Ala	Ile 85	Tyr	Tyr	Сув	Ala	Ala 90	Asp	Ser	Pro	Сув	Tyr 95	Met
40	Pro	Thr	Met	Pro 100	Ala	Pro	Pro	Ile	Arg 105	Asp	Ser	Phe	Gly	Trp 110	Asp	Asp
10	Phe	Gly	Gln 115	Gly	Thr	Gln	Val	Thr 120	Val	Ser	Ser					
45	(2) INFO	RMATI	ON F	FOR S	SEQ I	D NO	D: 29	ō:								
50	(i)	(B) (C)	LEN TYP STF	IGTH: PE: 8 RANDI	: 119 amino EDNES	am:	ino a id sing:	acid	5							
	(ii)	MOLE	CULE	TY	PE: p	prote	∍in									
55	(vii)	IMME (B)	CLC	TE SO	OURCI came	E: ∋1 "ì	neav	y cha	ain :	immuı	noglo	obul:	in" '	V-re	gion	(19)
	(xi)	SEQU	ENCE	DES	SCRI	PTIO	N: S	EQ II	ON C	: 25	:					
6 0	Gly 1	Gly	Ser	Val	Gln 5	Ala	Gly	Gly	Ser	Leu 10	Arg	Leu	Ser	Cys	Ala 15	Ala
65	Ser	Asp	Tyr	Thr 20	Ile	Thr	Asp	Tyr	Cys 25	Met	Ala	Trp	Phe	Arg 30	Gln	Ala
U.J	Pro	Gly	Lys 35	Glu	Arg	Glu	Leu	Val 40	Ala	Ala	Ile	Gln	Val 45	Val	Arg	Ser

		m b		T	m)	*		77.	3	C	11. 7	T	c1	7 ~~~	Dho	mh se
	Asp	50	Arg	Leu	Thr	Asp	Tyr 55	Ala	Asp	ser	vai	60	GIY	Arg	Phe	THE
5	Ile 65	Ser	Gln	Gly	Asn	Thr 70	Lys	Asn	Thr	Val	Asn 75	Leu	Gln	Met	Asn	Ser 80
	Leu	Thr	Pro	Glu	Asp 85	Thr	Ala	Ile	Tyr	Ser 90	Cys	Ala	Ala	Thr	Ser 95	Ser
10	Phe	Tyr	Trp	Tyr 100	Cys	Thr	Thr	Ala	Pro 105	Tyr	Asn	Val	Trp	Gly 110	Gln	Gly
15	Thr	Gln	Val 115	Thr	Val	Ser	Ser									
	(2) INFO	የ አልተ	ו מסו	7OR 9	SFO T	א ח	n. 26	5 •								
					_											
20	(1)	(B)	LENCE LENCE TYPE STE TO	NGTH: PE: 8 RANDE	: 117 amino EDNES	7 am: 5 ac: 5S: 9	ino a id sing:	acids	3							
25	(ii)	MOLE	ECULI	E TY	PE: 1	prote	∍in									
	(vii)						neavy	y cha	ain :	immu	noglo	obul:	in" '	V-re	gion	(20)
30	(xi)	SEQ	JENCI	E DES	CRI	PTIO	N: S1	EQ II	ON C	26:	:					
	Gly 1	Gly	Ser	Val	Gln 5	Val	Gly	Gly	Ser	Leu 10	Arg	Leu	Ser	Cys	Val 15	Ala
35	Ser	Thr	His	Thr 20	Asp	Ser	Ser	Thr	Cys 25	Ile	Gly	Trp	Phe	Arg 30	Gln	Ala
40	Pro	Gly	Lys 35	Glu	Arg	Glu	Gly	V al 4 0	Ala	Ser	Ile	Tyr	Phe 45	Gly	Asp	Gly
10	Gly	Thr 50	Asn	Tyr	Arg	Asp	Ser 55	Val	Lys	Gly	Arg	Phe 60	Thr	Ile	Ser	Gln
45	Leu 65	Asn	Ala	Gln	Asn	Thr 70	Val	Tyr	Leu	Gln	Met 75	Asn	Ser	Leu	Lys	Pro 80
	Glu	Asp	Ser	Ala	Met 85	Tyr	Tyr	Cys	Ala	Ile 90	Thr	Glu	Ile	Glu	Trp 95	Tyr
50	Gly	Cys	Asn	Leu 100	Arg	Thr	Thr	Phe	Thr 105	Arg	Trp	Gly	Gln	Gly 110	Thr	Gln
55	Val	Thr	Val 115	Ser	Ser											
	(2) INFO	RMAT:	ION I	FOR :	SEQ :	ID N); 2°	7:								
60	(i)	(B (C	UENCI) LEI) TYI) STI) TO	NGTH PE: 6 RANDI	: 12 amin EDNE:	5 am 5 ac SS:	ino a id sing	acid	S							
65	(ii)	MOL	ECUL	E TY	PE: 1	prot	ein									

(ii) MOLECULE TYPE: protein

	(vii)			re so one:			neavy	y cha	ain	immuı	noglo	obul:	in" '	/-re	gion	(21)
5	(xi)	SEQ	JENCI	E DES	SCRII	OITS	۷: SI	EQ II	OMC	27	:					
.,	Gly 1	Gly	Ser	Val	Gln 5	Val	Gly	Gly	Ser	Leu 10	Lys	Leu	Ser	Cys	Lys 15	Ile
10	Ser	Gly	Gly	Thr 20	Pro	Asp	Arg	Val	Pro 25	Lys	ser	Leu	Ala	Trp 30	Phe	Arg
	Gln	Ala	Pro 35	Glu	Lys	Glu	Arg	Glu 40	Gly	Ile	Ala	Val	Leu 45	Ser	Thr	Lys
15	Asp	Gly 50	Lys	Thr	Phe	Tyr	Ala 55	Asp	Ser	Val	Lys	Gly 60	Arg	Phe	Thr	Ile
20	Phe 65	Leu	Asp	Asn	Asp	Lys 70	Thr	Thr	Phe	Ser	Leu 75	Gln	Leu	Asp	Arg	Leu 80
	Asn	Pro	Glu	Asp	Thr 85	Ala	Asp	Tyr	Tyr	Cys 90	Ala	Ala	Asn	Gln	Leu 95	Ala
2 5	Gly	Gly	Trp	Tyr 100	Leu	Asp	Pro	Asn	Tyr 105	Trp	Leu	Ser	Val	Gly 110	Ala	Tyr
	Ala	Ile	Trp 115	Gly	Gln	Gly	Thr	His 120	Val	Thr	Val	Ser	Ser 12 5			
30	(2) INFO	RMATI	ION I	FOR S	SEQ I	ID NO	D: 28	3:								
35	(i)	(B)	LEI TYI STI	E CHA NGTH: PE: & RANDI POLOC	: 129 amino EDNES	ami aci	ino a id singl	cids	5							
4 0	(ii)	MOLI	ECULI	E TYI	PE: 1	prote	∍in									
	(vii)			TE SO ONE:			neavy	, cha	ain i	immuı	noglo	obul:	in" V	/-re	gion	(24)
45	(xi)	SEQU	JENCI	E DES	SCRII	PTIO	N: SI	EQ II	ON C	28:	:					
	Gly 1	Gly	Ser	Val	Gln 5	Ala	Gly	Gly	Ser	Leu 10	Arg	Leu	Ser	Cys	Asn 15	Val
50	Ser	Gly	Ser	Pro 20	Ser	Ser	Thr	Tyr	Cys 25	Leu	Gly	Trp	Phe	Arg 30	Gln	Ala
55	Pro	Gly	Lys 35	Glu	Arg	Glu	Gly	Val 40	Thr	Ala	Ile	Asn	Thr 45	Asp	Gly	Ser
	Val	Ile 50	Tyr	Ala	Ala	Asp	Ser 55	Val	Lys	Gly	Arg	Phe 60	Thr	Ile	Ser	Glm
60	Asp 65	Thr	Ala	Lys	Lys	Thr 70	Val	Tyr	Leu	Gln	Met 75	Asn	Asn	Leu	Gln	Pro 80
	Glu	Asp	Thr	Ala	Thr 85	Tyr	Tyr	Cys	Ala	Ala 90	Arg	Leu	Thr	Glu	Met 95	Gly
65	Ala	Cys	Asp	Ala 100	Arg	Trp	Ala	Thr	Leu 105	Ala	Thr	Arg	Thr	Phe	Ala	Tyr

Asn Tyr Trp Gly Arg Gly Thr Gln Val Thr Val Ser Ser

(2) INFORMATION FOR SEQ ID NO: 29: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 129 amino acids (B) TYPE: amino acid 10 (C) STRANDEDNESS: single (D) TOPOLOGY: linear (ii) MOLECULE TYPE: protein 15 (vii) IMMEDIATE SOURCE: (B) CLONE: camel "heavy chain immunoglobulin" V-region (25) (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 29: 20 Gly Gly Ser Val Gln Thr Gly Gly Ser Leu Arg Leu Ser Cys Glu Ile Ser Gly Leu Thr Phe Asp Asp Ser Asp Val Gly Trp Tyr Arg Gln Ala 25 Pro Gly Asp Glu Cys Lys Leu Val Ser Gly Ile Leu Ser Asp Gly Thr Pro Tyr Thr Lys Ser Gly Asp Tyr Ala Glu Ser Val Arg Gly Arg Val 30 Thr Ile Ser Arg Asp Asn Ala Lys Asn Met Ile Tyr Leu Gln Met Asn 75 35 Asp Leu Lys Pro Glu Asp Thr Ala Met Tyr Tyr Cys Ala Val Asp Gly Trp Thr Arg Lys Glu Gly Gly Ile Gly Leu Pro Trp Ser Val Gln Cys 40 Glu Asp Gly Tyr Asn Tyr Trp Gly Gln Gly Thr Gln Val Thr Val Ser 120 Ser 45 (2) INFORMATION FOR SEQ ID NO: 30: (i) SEQUENCE CHARACTERISTICS: 50 (A) LENGTH: 111 amino acids (B) TYPE: amino acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 55 (ii) MOLECULE TYPE: protein (vii) IMMEDIATE SOURCE: (B) CLONE: camel "heavy chain immunoglobulin" V-region (27) 60 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 30: Gly Gly Ser Val Gln Ala Gly Gly Ser Leu Arg Leu Ser Cys Ala Ser Ser Ser Lys Tyr Met Pro Cys Thr Tyr Asp Met Thr Trp Tyr Arg Gln 20 25 3065

	Ala	Pro	Gly 35	Lys	Glu	Arg	Glu	Phe 40	Val	Ser	Ser	Ile	Asn 45	Ile	Asp	Gly
5	Lys	Thr 50	Thr	Tyr	Ala	Asp	ser 55	Val	Lys	Gly	Arg	Phe 60	Thr	Ile	Ser	Gln
	Asp 65	Ser	Ala	Lys	Asn	Thr 70	Val	Tyr	Leu	Gln	Met 75	Asn	Ser	Leu	Lys	Pro 80
10	Glu	Asp	Thr	Ala	Met 85	Tyr	Tyr	Cys	Lys	Ile 90	Asp	Ser	Tyr	Pro	Cys 95	His
15	Leu	Leu	Asp	Val 100	Trp	Gly	Gln	Gly	Thr 105	Gln	Val	Thr	Val	Ser 110	Ser	
	(2) INFO	TAMS	ON I	FOR S	SEQ I	D NO	o: 3:	L:								
20	(i)	(A) (B) (C)	LEI TYI STI	E CHANGTH: PE: A RANDI	: 112 amino EDNES	2 ami 5 aci 5S: s	ino a id singl	cid	5							
25	(ii)	MOLI	CULI	TY!	PE: I	prote	∍in									
·	(vii)						neavy	/ cha	ain :	immu	noglo	bul:	in" '	/-re	gion	(29)
30	(xi)	SEQU	JENCI	E DES	SCRI	OITS	N: SI	EQ II	NO:	: 31	:					
		Gly	Ser	Val		Ala	Gly	Gly	Ser		Arg	Leu	Ser	Cys	Val 15	Ala
	1		_		5					10						
35	_	Gly	Phe	Asn 20	_	Glu	Thr	Ser	Arg 25		Ala	Trp	Tyr	Arg 30		Thr
35 40	Ser	_			Phe				25	Met			-	30	Gln	
	Ser Pro	Gly	Asn 35	20	Phe Cys	Glu	Leu	Val 40	25 Ser	Met	Ile	Tyr	Ser 45	30 Asp	Gln Gly	Lys
	Ser Pro Thr Asn 65	Gly Tyr 50 Ala	Asn 35 Tyr Lys	20 Val Val Asn	Phe Cys Asp	Glu Arg Leu 70	Leu Met 55 Tyr	Val 40 Lys Leu	ser Gly	Met Ser Arg	Ile Phe Ser 75	Tyr Thr 60	Ser 45 Ile Leu	30 Asp Ser Lys	Gln Gly Arg	Lys Glu Glu 80
40	Ser Pro Thr Asn 65	Gly Tyr 50 Ala	Asn 35 Tyr Lys	20 Val Val	Phe Cys Asp	Glu Arg Leu 70	Leu Met 55 Tyr	Val 40 Lys Leu	ser Gly	Met Ser Arg	Ile Phe Ser 75	Tyr Thr 60	Ser 45 Ile Leu	30 Asp Ser Lys	Gln Gly Arg	Lys Glu Glu 80
40	Ser Pro Thr Asn 65 Asp	Gly Tyr 50 Ala	Asn 35 Tyr Lys Ala	20 Val Val Asn	Phe Cys Asp Thr	Glu Arg Leu 70 Tyr	Leu Met 55 Tyr Cys	Val 40 Lys Leu Ala	25 Ser Gly Gln Pro	Met Ser Arg Leu Val	Ile Phe Ser 75 Glu	Tyr Thr 60 Gly Tyr	Ser 45 Ile Leu Pro	30 Asp Ser Lys	Gln Gly Arg Pro	Lys Glu Glu 80 Asp
40 45	Ser Pro Thr Asn 65 Asp Met	Gly Tyr 50 Ala Thr Cys	Asn 35 Tyr Lys Ala Ser	Val Val Asn Met Arg 100	Phe Cys Asp Thr Tyr 85 Tyr	Glu Arg Leu 70 Tyr Gly	Leu Met 55 Tyr Cys Asp	Val 40 Lys Leu Ala Pro	25 Ser Gly Gln Pro	Met Ser Arg Leu Val	Ile Phe Ser 75 Glu	Tyr Thr 60 Gly Tyr	Ser 45 Ile Leu Pro	30 Asp Ser Lys Ile Val	Gln Gly Arg Pro	Lys Glu Glu 80 Asp
40 45 50	Ser Pro Thr Asn 65 Asp Met	Gly Tyr 50 Ala Thr Cys RMAT: SEQUENCE (A (B) (C)	Asn 35 Tyr Lys Ala Ser ION UENCE LEE TYE	Val Val Asn Met Arg	Phe Cys Asp Thr Tyr 85 Tyr SEQ ARAC: 410 nucle	Glu Arg Leu 70 Tyr Gly ID No	Leu Met 55 Tyr Cys Asp O: 3: se peacid sing:	Val 40 Lys Leu Ala Pro	25 Ser Gly Gln Pro	Met Ser Arg Leu Val	Ile Phe Ser 75 Glu	Tyr Thr 60 Gly Tyr	Ser 45 Ile Leu Pro	30 Asp Ser Lys Ile Val	Gln Gly Arg Pro	Lys Glu Glu 80 Asp
40 45 50	Ser Pro Thr Asn 65 Asp Met	Gly Tyr 50 Ala Thr Cys SEQUENCE (A (B (C)	Asn 35 Tyr Lys Ala Ser ION UENCE UENCE TY:	Val Val Asn Met Arg 100 FOR SECHARISTER SE	Phe Cys Asp Thr Tyr 85 Tyr SEQ ARACC ARACC CHUCLE	Glu Arg Leu 70 Tyr Gly ID No FERIS 6 bas eic s eic s eic s eic s eic s	Leu Met 55 Tyr Cys Asp STIC: see pacid sing: ar	Val 40 Lys Leu Ala Pro	Ser Gly Gln Pro Gly 105	Met Ser Arg Leu Val	Ile Phe Ser 75 Glu	Tyr Thr 60 Gly Tyr	Ser 45 Ile Leu Pro	30 Asp Ser Lys Ile Val	Gln Gly Arg Pro	Lys Glu Glu 80 Asp

		(ix	(2	ATURI A) NI B) LO	AME/I			408									
5		(xi) SE	QUEN	CE DI	ESCR	IPTIC	ЭИ:	SEQ :	ID NO	o: 32	2:					
4.0										GGC Gly 10							48
10										AAC Asn							96
15										TGC Cys							144
20										GAA Glu							192
25										ACG Thr							240
3 0										TAT Tyr 90							288
50										GGA Gly							336
3 5										TCA Ser							384
40				GAC Asp				TAA'	TAGA	ATT (C						416
45	(2)		_	TION SEQUI						:							
50			(1	A) L1 B) T1 D) T0	YPE:	ami	no a	cid	aci	ds							
50		(ii) MO	LECU	LE T	YPE:	pro	tein									
				_					_	ID N			_		_		
5 5	Gln 1	Val	Lys	Leu	Leu 5	Glu	Ser	Gly	Gly	Gly 10	Ser	Val	Gln	Ala	Gly 15	Gly	
60	Ser	Leu	Thr	Leu 20	Ser	Cys	Val	Tyr	Thr 25	Asn	Asp	Thr	Gly	Thr 30	Met	Gly	
υυ	Trp	Phe	Arg 35	Gln	Ala	Pro	Gly	Lys 40	Glu	Cys	Glu	Arg	Val 45	Ala	His	Ile	
65	Thr	Pro	Asp	Gly	Met	Thr	Phe	Ile	Asp	Glu	Pro	Val	Lys	Gly	Arg	Phe	

	Thr 65	Ile	Ser	Arg	Asp	Asn 70	Ala	Gln	Lys	Thr	Leu 75	Ser	Leu	Arg	Met	Asn 80	
5	Ser	Leu	Arg	Pro	Glu 85	Asp	Thr	Ala	Val	Туг 90	Tyr	Cys	Ala	Ala	Asp 95	Trp	
	Lys	Tyr	Trp	Thr 100	Cys	Gly	Ala	Gln	Thr 105	Gly	Gly	Tyr	Phe	Gly 110	Gln	Trp	
10	Gly	Gln	Gly 115	Ala	Gln	Val	Thr	Val 120	Ser	Ser	Leu	Ala	Ser 125	Tyr	Pro	Tyr	
15	Asp	Val 130	Pro	Asp	Tyr	Gly	Se r 135										
	(2)			rion Quenc		_										·	
20		\- /	() (I	A) LE B) TY C) SI O) TO	ength (Pe : [rani	nucl	13 ba leic ESS:	acio sino	air:	3							
25		(ii)	MOI	LECUI	LE TY	PE:	DNA	(ger	nomio	=)							
30	•	(vii)		MEDIA B) CI		can	nel '		-			nogi B09		lin"	V-re	egion	followed
, ,		(ix)	(1	ATURE A) NA B) LO	AME/I	KEY:		135									
			•	-,													
35		(xi)		QUENC			PTIC	on: S	SEQ :	ID NO): 3 ⁴	4 :					
		GTG	SE(•	CTC	escri Gag	TCT	GGA	GGA	GGC	TCG	GTG					48
35 40	Gln 1 TCT	GTG Val	SE(AAA Lys AGA	OUENC CTG	CTC Leu 5	GAG Glu TGT	TCT Ser	GGA Gly GTC	GGA Gly	GGC Gly 10 GGA	TCG Ser	GTG Val	Gln TTT	Thr	Gly 15 ACC	Gly AGT	4 8
	Gln 1 TCT Ser	GTG Val CTG Leu	AAA Lys AGA Arg	CTC Leu Leu	CTC Leu 5 TCC Ser	GAG Glu TGT Cys	TCT Ser GCA Ala	GGA Gly GTC Val	GGA Gly TCT Ser 25	GGC Gly 10 GGA Gly	TCG Ser TTC Phe	GTG Val TCC Ser	Gln TTT Phe CGT	Thr AGT Ser 30 GAG	Gly 15 ACC Thr	Gly AGT ser	
40	Gln 1 TCT Ser TGT Cys	GTG Val CTG Leu ATG Met	AAA Lys AGA Arg GCC Ala 35	CTG Leu CTC Leu 20	CTC Leu 5 TCC Ser TTC Phe	GAG Glu TGT Cys CGC Arg	TCT Ser GCA Ala CAG Gln	GGA Gly GTC Val GCT Ala 40 GGT	GGA Gly TCT Ser 25 TCA Ser	GGC Gly 10 GGA Gly GGA Gly	TCG Ser TTC Phe AAG Lys	GTG Val TCC Ser CAG Gln	Gln TTT Phe CGT Arg 45	Thr AGT Ser 30 GAG Glu ACA	Gly 15 ACC Thr GGG Gly	Gly AGT Ser GTC Val	96
40 45	Gln 1 TCT Ser TGT Cys GCA Ala	GTG Val CTG Leu ATG Met GCC Ala 50	AAA Lys AGA Arg GCC Ala 35 ATT Ile	CTG Leu CTC Leu 20 TGG Trp	CTC Leu 5 TCC Ser TTC Phe AGT Ser	GAG Glu TGT Cys CGC Arg	TCT Ser GCA Ala CAG Gln GGT Gly 55	GGA Gly GTC Val GCT Ala 40 GGT Gly	GGA Gly TCT Ser 25 TCA Ser AGG Arg	GGC Gly 10 GGA Gly GGA Gly ACA Thr	TCG Ser TTC Phe AAG Lys TAC Tyr	GTG Val TCC Ser CAG Gln TAC Tyr 60 CAA	Gln TTT Phe CGT Arg 45 AAC Asn	AGT Ser 30 GAG Glu ACA Thr	Gly 15 ACC Thr GGG Gly TAT Tyr	AGT SET GTC Val	96 144
40 45 50	Gln 1 TCT Ser TGT Cys GCA Ala GCC Ala 65	GTG Val CTG Leu ATG Met GCC Ala 50 GAG Glu	SEC AAA Lys AGA Arg GCC Ala 35 ATT Ile TCC Ser	CTG Leu CTC Leu 20 TGG Trp AAT Asn	CTC Leu 5 TCC Ser TTC Phe AGT Ser AAG Lys	GAG GLU TGT Cys CGC Arg GGC Gly GGC Gly 70 GAT	TCT Ser GCA Ala CAG Gln GGT Gly 55 CGA Arg	GGA Gly GTC Val GCT Ala 40 GGT Gly TTC Phe	GGA Gly TCT Ser 25 TCA Ser AGG Arg	GGC Gly 10 GGA Gly GGA Thr	TCG Ser TTC Phe AAG Lys TAC Tyr TCC Ser 75	GTG Val TCC Ser CAG Gln TAC Tyr 60 CAA Gln	Gln TTT Phe CGT Arg 45 AAC Asn GAC Asp	AGT Ser 30 GAG Glu ACA Thr AAC Asn	Gly 15 ACC Thr GGG Gly TAT Tyr GCC Ala	AGT Ser GTC Val GTC Val AAG Lys 80 GCT	96 144 192
40 45	Gln TCT Ser TGT Cys GCA Ala GCC Ala 65 ACC Thr	GTG Val CTG Leu ATG Met GCC Ala 50 GAG Glu ACG Thr	SEC AAA Lys AGA Arg GCC Ala 35 ATT Ile TCC Ser GTA Val	CTG Leu 20 TGG Trp AAT Asn GTG Val	CTC Leu 5 TCC Ser TTC Phe AGT Ser AAG Lys CTT Leu 85 GCG	GAG Glu TGT Cys CGC Arg GGC Gly GGT Asp	TCT Ser GCA Ala CAG Gln GGT Gly 55 CGA Arg ATG Met	GGA Gly GTC Val GCT Ala 40 GGT Gly TTC Phe AAC ASN	GGA Gly TCT Ser 25 TCA Ser AGG Arg GCC Ala	GGC Gly 10 GGA Gly ACA Thr ATC Ile CTA Leu 90	TCG Ser TTC Phe AAG Lys TAC Tyr TCC Ser 75 ACC Thr	GTG Val TCC Ser CAG Gln TAC Tyr 60 CAA Gln CCT Pro	Gln TTT Phe CGT Arg 45 AAC Asn GAC Asp GAA Glu CCT	AGT Ser 30 GAG Glu ACA Thr AAC Asn GAC Asp GGC	Gly ACC Thr GGG Gly TAT Tyr GCC Ala ACG Thr 95	AGT Ser GTC Val GTC Val AAG Lys 80 GCT Ala	96 144 192 240

	GTC Val	TCC Ser 130	TCA Ser	CTA Leu	GCT Ala	AGT Ser	TAC Tyr 135	CCG Pro	TAC Tyr	GAC Asp	GTT Val	CCG Pro 140	GAC Asp	TAC Tyr	GGT Gly	TCT Ser	. 43:	2
5	TAAT	TAGAA	ATT (C													443	3
	145																	
10	(2)	INFO				-												
15		((/ (I	SEQUIA) LIB) TY	ENGTI PE:	H: 14 amir	14 an	nino cid										
		(ii)	MOI	LECUI	LE TY	PE:	prot	ein										
20		(xi)	SE	QUENC	CE DE	ESCRI	PTIC	on: s	SEQ 3	ID NO	35	5:						
	Gln 1	Val	Lys	Leu	Leu 5	Glu	Ser	Gly	Gly	Gly 10	Ser	Val	Gln	Thr	Gly 15	Gly		
25	Ser	Leu	Arg	Leu 20	Ser	Cys	Ala	Val	Ser 25	Gly	Phe	Ser	Phe	Ser 30	Thr	Ser		
	Cys	Met	Ala 35	Trp	Phe	Arg	Gln	Ala 40	Ser	Gly	Lys	Gln	Arg 45	Glu	Gly	Val		
30	Ala	Ala 50	Ile	Asn	Ser	Gly	Gly 55	Gly	Arg	Thr	Tyr	Tyr 60	Asn	Thr	Tyr	Val		
35	Ala 65	Glu	Ser	Val-	- Lys	Gly 70	Arg	Phe	Ala	Ile	Ser 75	Gln	Asp	Asn	Ala	Eys 80		
	Thr	Thr	Val	Tyr	Leu 85	Asp	Met	Asn	Asn	Leu 90	Thr	Pro	Glu	Asp	Thr 95	Ala		
40	Thr	Tyr	Tyr	Cys 100	Ala	Ala	Val	Pro	Ala 105	His	Leu	Gly	Pro	Gly 110	Ala	Ile		
	Leu	Asp	Leu 115	Lys	Lys	Tyr	Lys	Tyr 120	Trp	Gly	Gln	Gly	Thr 125	Gln	Val	Thr		
45	Val	Ser 130	Ser	Leu	Ala	Ser	Tyr 135	Pro	Tyr	Asp	Val	Pro 140	Asp	Tyr	Gly	Ser		
50	(2)	INFO) SE	QUEN	CE CI	HARAG	CTER:	ISTIC	cs:									
55			(A) LI B) Ti C) Si D) To	YPE: TRANI	nuc: DEDNI	leic ESS:	acio	d	5								
		(ii)) MO	LECU:	LE T	YPE:	DNA	(gei	nomi	c)								
60		(vii)) IM	MEDI	ATE :	SOUR(CE: mel 1	heav	y ch	ain		nogle pB24		in"	V-re	gion	followed	
65		(ix)	· (.	ATUR A) N B) L	AME/			441										
		(xi) SE	QUEN	CE D	ESCR	IPTI	: NC	SEQ	ID N	0: 3	6:						

	CAG Gln 1	GTG Val	AAA Lys	CTG Leu	CTC Leu 5	GAG Glu	TCT Ser	GGG Gly	GGA Gly	GGG Gly 10	TCG Ser	GTG Val	CAG Gln	GCT Ala	GGA Gly 15	GGG GGG	4 8	
5												CCC Pro					96	
10	TGC Cys	CTG Leu	GGC Gly 35	TGG Trp	TTC Phe	CGC Arg	CAG Gln	GCT Ala 40	CCA Pro	GGG Gly	AAG Lys	GAG Glu	CGT Arg 45	GAG Glu	GGG Gly	GTC Val	144	
15												GCA Ala 60					192	
20	AAG Lys 65	GGC Gly	C GA A rg	TTC Phe	ACC Thr	ATC Ile 70	TCC Ser	CAA Gln	GAC Asp	ACC Thr	GCC Ala 75	AAG Lys	AAA Lys	ACG Thr	GTA Val	TAT Tyr 80	240	
												GCC Ala					28 8	
25	GCG Ala											GCG Ala					33 6	
30												GGC Gly					384	
35	GTC Val	ACC Thr 130	GTC Val	TCC Ser	TCA Ser	CTA Leu	GCT Ala 135	AGT Ser	TAC Tyr	CCG Pro	TAC Tyr	GAC Asp 140	GTT Val	CCG Pro	GAC Asp	TAC Tyr	432	
40	GGT Gly 145		TAAT	'AGA <i>I</i>	ATT C	;											449	
	(2)	INFO	RMAT	NOI	FOR	SEQ	ID N	10: 3	37:									
45		•	(2 (E	A) LE B) TY	NGTI PE:	: 14 amir	RACTE 16 an no ac line	nino cid										
50		(ii)	MOI	LECUI	E TY	PE:	prot	ein										
50		(xi)	SEÇ	QUENC	E DE	SCRI	PTIC	on: s	SEQ 1	D NC): 37	7:						
55	Gln 1	Val	Lys	Leu	Leu 5	Glu	Ser	Gly	Gly	Gly 10	Ser	Val	Gln	Ala	Gly 15	Gly	•	
	Ser	Leu	Arg	Leu 20	Ser	Cys	Asn	Val	Ser 25	Gly	Ser	Pro	Ser	Ser 30	Thr	Tyr		
60	Cys	Leu	Gly 35	Trp	Phe	Arg	Gln	Ala 40	Pro	Gly	Lys	Glu	Arg 45	Glu	Gly	Val		
	Thr	Ala 50	Ile	Asn	Thr	Asp	Gly 55	Ser	Val	Ile	Tyr	Ala 60	Ala	Asp	Ser	Val		
65	Lys 65	Gly	Arg	Phe	Thr	Ile 70	Ser	Gln	Asp	Thr	Ala 75	Lys	Lys	Thr	Val	Tyr 80		

	Leu	Gln	Met	Asn	Asn 85	Leu	Gln	Pro	Glu	Asp 90	Thr	Ala	Thr	Tyr	Tyr 95	Cys .		
5	Ala	Ala	Arg	Leu 100	Thr	Glu	Met	Gly	Ala 105	Сув	Asp	Ala	Arg	Trp 110	Ala	Thr		
	Leu	Ala	Thr 115	Arg	Thr	Phe	Ala	Tyr 120	Asn	Tyr	Trp	Gly	Arg 125	Gly	Thr	Gln		
10	Val	Thr 130	Val	Ser	Ser	Leu	Ala 135	Ser	Tyr	Pro	туг	Asp 140	Val	Pro	Asp	Tyr		
15	Gly 145	Ser																
	(2)	INF	ORMA'	TION	FOR	SEQ	ID I	NO: .	38:									
20		(i	() ()	QUENCA) LI B) TI C) SI	ENGTI YPE: IRANI	nuc DEDNI	19 ba leic ESS:	ase ; acio sino	pair: d	5								
25		(ii) MO:	LECU	LE T	YPE:	DNA	(ge	nomi	c)								
•		(vii	•	MEDIA B) C				gure	6									
30		(xi) SE	QUEN	CE D	ESCR	IPTI(ON:	SEQ	ID N	o: 3	B:						
	AAT	TTAG	CGG	CCGC	CCAG	GT G	AAAC'	TGCT	C GA	GTAA	GTGA	CTA	AGGT	CAC	CGTC	TCCTC	4	60
35	GAA	CAAA	AAC '	TCAT	CTCA(GA A	GAGG	ATCT	G AA	TTAA'	rgag	AAT	CAT	CAA i	ACGG	TGATA		119
	(2)	INF	ORMA	TION	FOR	SEQ	ID !	NO:	39:									
40		(i	· (QUEN A) L B) T C) S D) T	ENGT YPE: TRAN	H: 1 nuc DEDN	20 b leic ESS:	ase aci sin	pair d	S								
45		(ii) мо	LECU	LE T	YPE:	DNA	(ge	nomi	c)								
		(Vii) IM	MEDI B) C	ATE LONE	SOUR : Se	CE: e fi	gure	6									
50		(xi) SE	QUEN	CE D	ESCR	IPTI	ON:	SEQ	ID N	0: 3	9:						
	AGC	TTAT	CAC	CGTT	TGAT	GA A	TTCT	CATT	A AT	TCAG	ATCC	TCT	TCTG	AGA	TGAG	TTTTT	3	60
55	TTC	TGAG	GAG	ACGG	TGAC	CT T	AGTC	ACTT	A CT	CGAG	CAGT	TTC	ACCT	GGG	CGGC	CGCTA	A	120
	(2)	INF	ORMA	TION	FOR	SEQ	ID	NO:	40:									
60		(i	· (QUEN A) L B) T C) S D) T	ENGT YPE: TRAN	H: 7 ami DEDN	ami no a ESS:	no a cid sin	cids									
65		/::) MC	ווסים זו	ייי יבו	vor.	220	+ain										

	<pre>(vii) IMMEDIATE SOURCE: (B) CLONE: See figure 6</pre>	
5	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 40:	
5	Ala Gln Val Lys Leu Glu 1 5	
10	(2) INFORMATION FOR SEQ ID NO: 41:	
15	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 16 amino acids (B) TYPE: amino acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 	
	(ii) MOLECULE TYPE: protein	
20	<pre>(vii) IMMEDIATE SOURCE: (B) CLONE: See figure 6</pre>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 41:	
2 5	Val Thr Val Ser Ser Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu Asn 1 5 10 15	
30	(2) INFORMATION FOR SEQ ID NO: 42:	
35	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 117 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 	
	(ii) MOLECULE TYPE: DNA (genomic)	
40	(vii) IMMEDIATE SOURCE: (B) CLONE: See figure 19	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 42:	
45	AATTTAGTCG CGACAGGTGA AACTGCTCGA GTAAGTGACT AAGGTCACCG TCTCCTCAGA ACAAAAACTC ATCTCAGAAG AGGATCTGAA TTAATGAGAA TTCATCTTAA GGTGATA	60 117
50	(2) INFORMATION FOR SEQ ID NO: 43:	
	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 117 base pairs (B) TYPE: nucleic acid	
55	(C) STRANDEDNESS: single (D) TOPOLOGY: linear	
	(ii) MOLECULE TYPE: DNA (genomic)	
60	(vii) IMMEDIATE SOURCE: (B) CLONE: See figure 19	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 43:	
65	AGCTTATCAC CTTAAGATGA ATTCTCATTA ATTCAGATCC TCTTCTGAGA TGAGTTTTTG	60
	THETELOGIAC NORTH TACTORETTA CHOCACONCH THEORETTE COCACHA	117

	(2) INFO	RMATION FOR SEQ 1D NO: 44:
5	(i)	SEQUENCE CHARACTERISTICS: (A) LENGTH: 6 amino acids (B) TYPE: amino acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear
10	(ii)	MOLECULE TYPE: protein
-0	(Vii)	IMMEDIATE SOURCE: (B) CLONE: See figure 19
15	(xi)	SEQUENCE DESCRIPTION: SEQ ID NO: 44:
	Arg 1	Gln Val Lys Leu Leu 5
20	(2) INFO	RMATION FOR SEQ ID NO: 45:
2 5	(i)	SEQUENCE CHARACTERISTICS: (A) LENGTH: 16 amino acids (B) TYPE: amino acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear
	(ii)	MOLECULE TYPE: protein
30	(vii)	IMMEDIATE SOURCE: (B) CLONE: See figure 19
	(xi)	SEQUENCE DESCRIPTION: SEQ ID NO: 45:
35	Val 1	Thr Val Ser Ser Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu Asn 5 10 15
40	(2) INFO	RMATION FOR SEQ ID NO: 46:
	(i)	SEQUENCE CHARACTERISTICS: (A) LENGTH: 4 amino acids (B) TYPE: amino acid
45		(C) STRANDEDNESS: single(D) TOPOLOGY: linear
	(ii)	MOLECULE TYPE: protein
50	(xi)	SEQUENCE DESCRIPTION: SEQ ID NO: 46:
	Gln 1	Val Lys Leu
55	(2) INFO	RMATION FOR SEQ ID NO: 47:
60	(i)	SEQUENCE CHARACTERISTICS: (A) LENGTH: 5 amino acids (B) TYPE: amino acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear
	(ii)	MOLECULE TYPE: protein
65		SEQUENCE DESCRIPTION: SEO ID NO: 47:

Val Thr Val Ser Ser

70

	1 5	•
5	(2) INFORMATION FOR SEQ ID NO: 48:	
10	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 21 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear	
	(ii) MOLECULE TYPE: DNA (genomic)	
15	(xi) SEQUENCE DESCRIPTION: SEQ ID	NO: 48:
	GTCACCGTCT CCTCATAATG A	21
20	(2) INFORMATION FOR SEQ ID NO: 49:	
25	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 20 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 	•
	(ii) MOLECULE TYPE: DNA (genomic)	
30	(xi) SEQUENCE DESCRIPTION: SEQ ID	NO: 49:
	AGCTTCATTA TGAGGAGACG	20
35	/2) THEODY MICH CO. TO NO. FO.	
33	(2) INFORMATION FOR SEQ ID NO: 50:	
40	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 34 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 	
	(ii) MOLECULE TYPE: DNA (genomic)	
45	(xi) SEQUENCE DESCRIPTION: SEQ ID	NO: 50:
	GTCACCGTCT CCTCATAATG ATCTTAAGGT GATA	34
50	(2) INFORMATION FOR SEQ ID NO: 51:	
55	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 33 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 	
	(ii) MOLECULE TYPE: DNA (genomic)	
60	(xi) SEQUENCE DESCRIPTION: SEQ ID	NO: 51:
	ACCTTATCAC CTTAACATCA TTATCACCAC ACC	33

65

	(2) INFORMATION FOR SEQ ID NO: 52:	
5	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 12 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 	
10	(ii) MOLECULE TYPE: DNA (genomic)	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 52:	
	AATTGCGGCC GC	12
15	(2) INFORMATION FOR SEQ ID NO: 53:	
20	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 16 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 	
25	(ii) MOLECULE TYPE: DNA (genomic)	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 53:	
	CATGCAGTCT TCGGGC	16
30	(2) INFORMATION FOR SEQ ID NO: 54:	
35	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 16 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 	
40	(ii) MOLECULE TYPE: DNA (genomic)	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 54:	
	TTAAGCCCGA AGACTG	16
45	(2) INFORMATION FOR SEQ ID NO: 55:	
50	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 44 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear	
55	(ii) MOLECULE TYPE: DNA (genomic)	
JJ	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 55:	
	TCACTGAATT CGGGATCATG AGGACTCTCC TTGTGAGCTC GCTT	44
60	(2) INFORMATION FOR SEQ ID NO: 56:	
65	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 48 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 	

72

	(ii) MOLECULE TYPE: DNA (genomic)	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 56:	
5	ATGTCACAAA GCTTAAGCAC GAAGACAGTC GACCGTGCGG CCGGAGAC	48
	(2) INFORMATION FOR SEQ ID NO: 57:	
10	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 44 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 	
15	(ii) MOLECULE TYPE: DNA (genomic)	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 57:	
20	CGCGTCCATG CAGTCCTCAG GTGGATCATC CCAGGTGAAA CTGC	44
	(2) INFORMATION FOR SEQ ID NO: 58:	
25	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 44 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single	
30	(D) TOPOLOGY: linear	
	(ii) MOLECULE TYPE: DNA (genomic)	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 58:	
3 5	TCGAGCAGTT TCACCTGGGA TGATCCACCT GAGGACTGCA TGGA	44
	(2) INFORMATION FOR SEQ ID NO: 59:	
40	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 15 amino acids (B) TYPE: amino acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 	
45	(ii) MOLECULE TYPE: protein	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 59:	
50	Ser Met Gln Ser Ser Gly Gly Ser Ser Gln Val Lys Leu Leu Glu 1 5 10 15	
55	(2) INFORMATION FOR SEQ ID NO: 60:	
60	(i) SEQUENCE CHARACTERISTICS:(A) LENGTH: 53 base pairs(B) TYPE: nucleic acid(C) STRANDEDNESS: single(D) TOPOLOGY: linear	
00	(ii) MOLECULE TYPE: DNA (genomic)	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 60:	
65	CATGGCCAGG TGAAACTGCT CGAGTAAGTG ACTAAGGTCA CCGTCTCCTC AGC	53

PCT/EP94/01442

	(2) INFORMATION FOR SEQ ID NO: 61:	
5	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 53 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 	
10	(ii) MOLECULE TYPE: DNA (genomic)	
10	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 61:	
	GGCCGCTGAG GAGACGGTGA CCTTAGTCAC TTACTCGAGC AGTTTCACCT GGC	53
15	(2) INFORMATION FOR SEQ ID NO: 62:	
2 0	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 6 amino acids (B) TYPE: amino acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 	
25	(ii) MOLECULE TYPE: protein	
4 J	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 62:	
3 0	Ser Ser Gly Gly Ser Ser 1 5	

CLAIMS

- 1. A process for the production of an antibody or a fragment or functionalized fragment thereof using a transformed lower eukaryotic host containing an expressible DNA sequence encoding the antibody or (functionalized) fragment thereof, wherein the antibody or (functionalized) fragment thereof is derived from a heavy chain immunoglobulin of *Camelidae* and is devoid of light chains, and wherein the lower eukaryotic host is a mould or a yeast.
- 2. A process according to claim 1, in which the mould belongs to the genera Aspergillus or Trichoderma.
 - 3. A process according to claim 1, in which the yeast belongs to the genera Saccharomyces, Kluyveromcyes, Hansenula, or Pichia.

15

30

5

- 4. A process according to claim 1, in which the heavy chain fragment at least contains the whole variable domain.
- 5. A process according to claim 1, in which the antibody or (functionalized)

 fragment thereof derived from a heavy chain immunoglobulin of Camelidae
 comprises a complementary determining region (CDR) different from the CDR
 belonging to the natural antibody ex Camelidae grafted on the framework of the
 variable domain of the heavy chain immunoglobulin ex Camelidae.
- 6. A process according to claim 1, in which the immunoglobulin to be produced is a catalytic antibody raised in *Camelidae*.
 - 7. A process according to claim 1, in which the functionalized antibody or fragment thereof comprises a fusion protein of both a heavy chain immunoglobulin from *Camelidae* or a fragment thereof and another polypeptide.

8. A process according to claim 1, in which the DNA sequence encodes a modified heavy chain immunoglobulin or (functionalized) fragment thereof derived from *Camelidae* and being devoid of light chains, and is made by random or directed mutagenesis or both.

5

10

- 9. A process according to claim 8, in which the resulting immunoglobulin or (functionalized) fragment thereof is modified such that
 - it is better adapted for production by the host cell, or
 - it is optimized for secretion by the lower eukaryotic host into the fermentation medium, or
 - its binding properties (k_{on} and k_{off}) are optimized, or
 - its catalytic activity is improved, or
 - it has acquired a metal chelating activity, or
 - its physical stability is improved.

15

20

- 10. A composition containing a product produced by a process as claimed in any one of claims 1-9.
- 11. New product obtainable by a process as claimed in any one of claims 1-9.
 - 12. A composition containing a new product as claimed in claim 11.

* * * * *

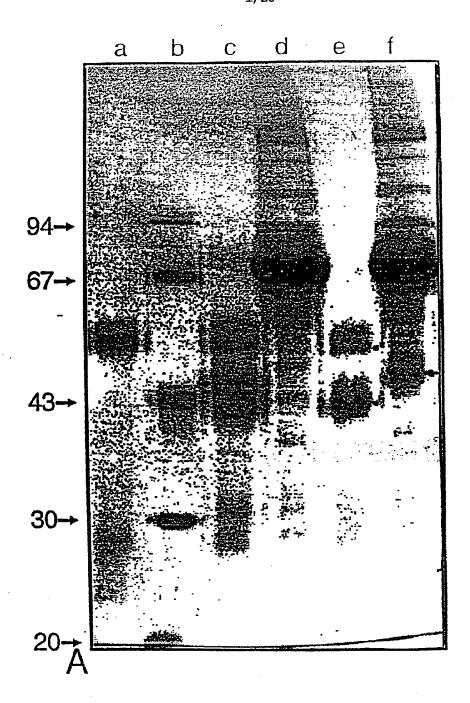
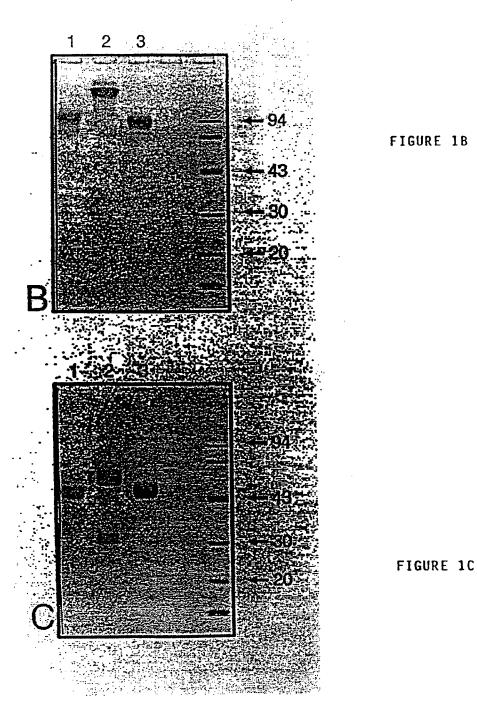
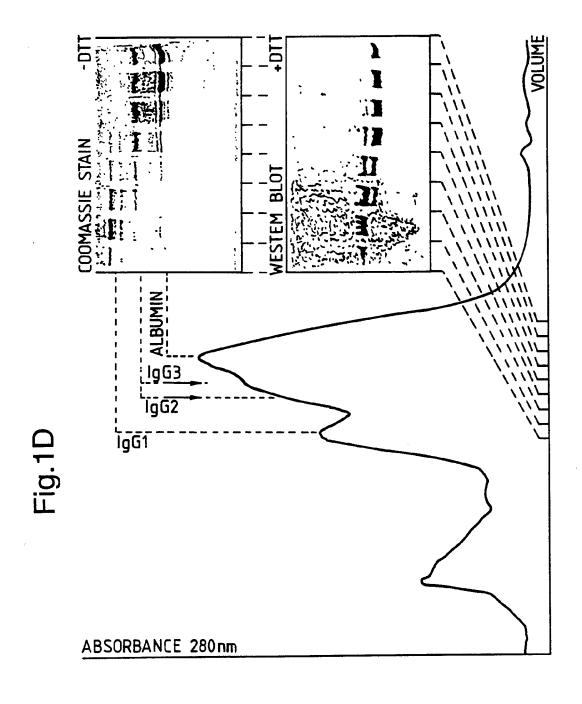


FIGURE 1A





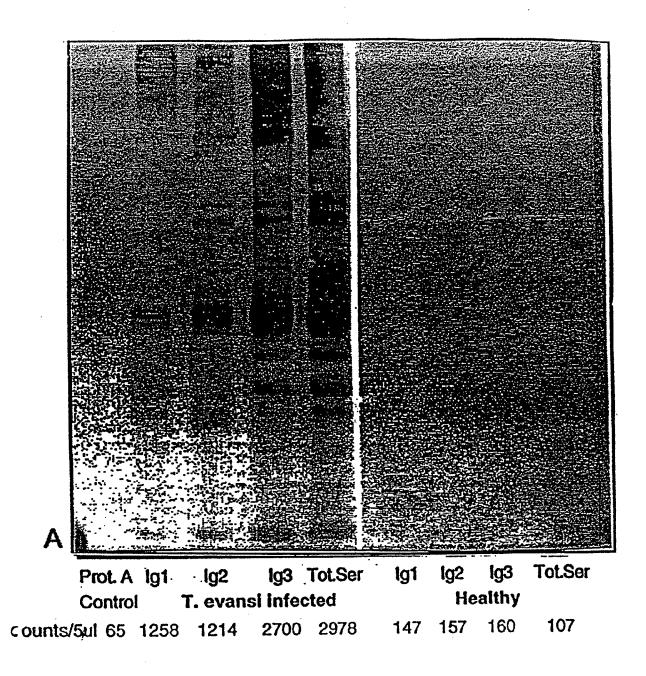


FIGURE 2A

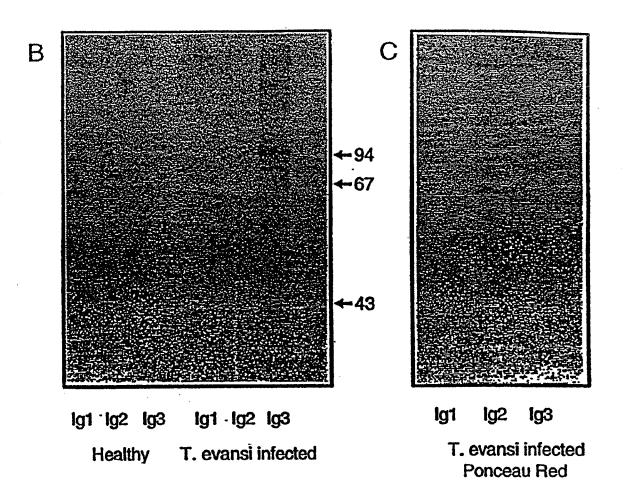


FIGURE 2B FIGURE 2C

Fig.3.	20			40		••••
EVQLVESGGG	LVQPGGSLRL	SCAASG	CDRI	WVRQA	PGKGLEWVS	CDR2
GG	SVQGGGSLRL	SCAISG	CDR1	WFREG	PGKEREGIA	CDR2
GG	SVQAGGSLRL	SCASSS	CDR1	WYRQA	PGKEREFVS	CDR2

70	80	90				
RFTIS	RDNSKNTLYL	OMNSLRAEDTAVY	YCAR	CDR3	WGQGTLVT	VSS
RFTIS	QDSTLKTMYL	LMNNLKPEDTGTY	YCAA	CDR3	WGQGTQVT	vss
RFTIS	QDSAKNTVYL	QMNSLKPEDTAMY	YCKI	CDR3	WGQGTQVT	vss

	cam	el V _H	hinge	C _H 2
7	WGQGT	QVT VSS	GTNEVCKCPKCP	APELPGG PSVFVFP
camel	WGQGT	QVT VSS	- EPKIPQPQPKPQPQP	•
			QPQPKPQP	
			KPEPECTCPKCP	APELLGG PSVFIFP
• • • • • •	huma	n C _H l	hinge	C _H 2
human g	amma 3	KVDKRV	ELKTPLGDTTHTCPRCP	•
			EPKCSDTPPPCPRCP	•
			EPKSCDTPPPCPRCP	APELLGG PSVFLFP
human g	amma 1	KVDKK-	AEPKSCDKTHTCPPCP	APELLGG PSVFLFP
human g	ramma 2	KVKVTV	ERKCCVECPPCP	APPVAG- PSVFLFP
human g	ramma 4	KVDKRV	ESKYGPPCPSCP	APEFLGG PSVFLFP

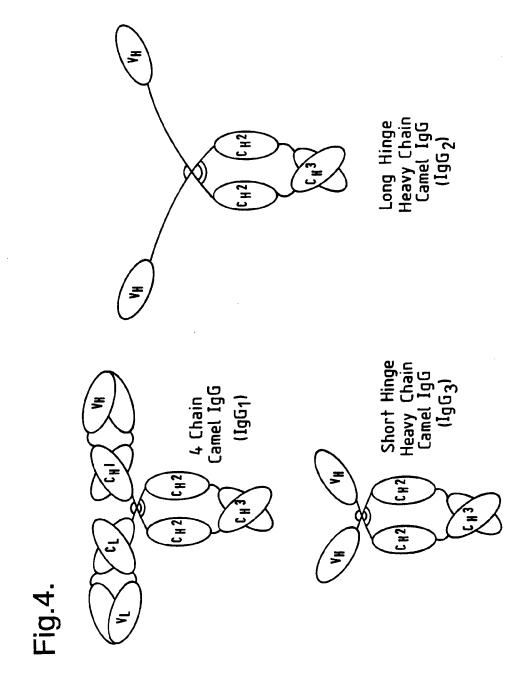


Fig.5A.

1				-+-	GC1		+				-+			+-						ACTC	60
	GT	CCY	CTT	TGA	.CGA	GC1	CVC	ACC	TCC	TCC	GAC	SCC)	CGI	CTC	SACC	rcc	:ፐለር	λGλ	CTC	TGAG	
	Q	v	K	L	L	E	S	G	G	G	s	v	Q	T	G	G	S	L	R	L	-
61		CTG																		GCT	120
-																				CCGA	
	s	С	A	v	s	G	F	s	F	s	Т	s	С	M	A	W	F	R	Q	A	-
121																				CTAC	180
121																				GATG	100
	s	G	ĸ	Q	R	E	G	v	A	A	I	N	s	G	G	G	R	T	Y	Y	-
101																				CAAG	240
101																				CITC	240
	N	T	Y	v	A	E	s	v	ĸ	G	R	F	A	I	s	Q	D	N	A	ĸ	-
242																				CTGT	200
241				-+-							+			-+-			+			CTGT + GACA	300
241		GTG		TAT.		ACT					+	GGG		-+-			+ ATG	CAT		GACA	300 -
	TGG T	T GGC	CCA' V	TAT.	AGA L	ACT D	ATA M	CTT N GGG	GTT N ACC	GGA L TGG	TTG	GGG P	ACT E TCT	TCT D	GTG T	CCG A GAA	T AAA	CAT Y GTA	AAT Y TAA	GACA C GTAC	_
	TGG	T GGC	CCA V GGT	TAT.	AGA L AGC	D CCA	ATA M	CTT N GGG	GTT N ACC	GGA L TGG	TTG	P CAT	E TCT	TCT D	T T	CCG A GAA	T AAA	CAT Y GTA	AAT Y TAA	GACA C	_
	T	T GGC	CCA V GGT	Y CCC.	AGA L AGC	ACT D CCA GGT	ATA M	CTT N GGG	GTT N ACC	GGA L TGG	T T CGC	GGG P CAT	E TCT	TGA	T T TTT	CCG A GAA CTT	T AAA	CAT Y GTA CAT	AAT Y TAA	GACA C GTAC	_
	TGGGGGGGA	T GGC: CCG	CCA V GGT CCA	TAT. Y CCC. GGG	AGA L AGC TCG A	ACT D CCA GGT H	CTT GAA L	CTT N GGG CCC	GTT N ACC TGG	TGG ACC	TTGCCCC	GGG P CAT	E TCT AGA	TGA ACT	T TTT TAAA L	CCG A GAA CTT	T AAA T TTT	CAT GTA CAT	AAT TAA ATT	GACA C GTAC CATG	_
301	TGGGGGA	GTG T GGC: CCG	CCA CCA	TAT. Y CCC. F	AGA L AGC TCG A	ACT D CCA GGT H	ATA M CTT GAA L Bst	GGG GGG GEII	ACC TGG	GGA L TGG ACC	TTG T CGC A	GGG P CAT	ACT E TCT AGA L	TGA ACT D	TTT AAA	GAA CTT K	T AAA TTTT K GTA	CAT Y GTA CAT Y	TAA ATT K	GACA C GTAC CATG Y TCCG	- 360 -
301	TGC CGC A	T GGC: CCG	CCA CCA CCA	TAT. Y CCC. GGG	AGA L AGC TOG A GAC	ACT D CCA GGT CCA GGT	M CTT GAA L Bst CCA	CTT N GGG GCC G	ACC TGG P	GGA TGG ACC G CTC	T C C C C C C C C C C C C C C C C C C C	GGG P CAT GTA I	E TCT AGA L	TGA ACT D	TTTT AAA	GAAA CTTT K	T AAA TTTT K GTA CAT	CAT Y CAT Y CGA	TAA ATT K CGT	GTAC GTAC CATG Y TCCG AGGC	- 360 -
301	TGC CGC A	GTG T GGC: CCG	CCA CCA CCA CCA	TAT. Y CCC. F	AGA L AGC TOG A GAC	ACT D CCA GGT CCA GGT	ATA M CTT GAA L Bst	CTT N GGG GCC G	ACC TGG P	GGA TGG ACC G CTC	TTG T CGC A	GGG P CAT	E TCT AGA L	TGA ACT D	TTT AAA	GAA CTT K	T AAA TTTT K GTA	CAT Y GTA CAT Y	TAA ATT K	GACA C GTAC CATG Y TCCG	- 360 -

Fig.5B.

1				-+-	GCT		+				-+			-+-			+			ACTC	60
•	GT	CCY	CTT	TGA	CGA	GCI	CAC	ACC	ccc	CTCC	:GAG	CCA	CGT	,CCC	ACC	ccc	CAC	AGA	CTG	TGAG	
	Q	v	к	L	L	Ε	s	G	G	G	s	v	Q	Α	G	G	s	L	T	L	-
										N	tyI coI										
61	TC	TTG	TGT	ATA	CAC	CVY	CGA	TAC	TGC	GAC	CAT	GGG	ATG	GTI	TCG	CCA	GGC	TCC	AGG	GAAA	120
91	λG	VVC	ACΛ	TAT	GTG	GTT	'GC'I	' ለ ፐር	λCC	CTG	ĠΊλ	CCC	TAC	CYV.	λGC	:GGT	.ccc	AGG	TCC	CTTT	100
	s	С	v	Y	T	N	D	T	G	T	М	G	W	F	R	Q	Α	P	G	K	-
•	Gλ	GTG	CGA	AAG	GGT	CGC	:GCA	TAT.	TAC	GCC	TGA	TGG	ТАТ	GAC	CTT	'CAT	TGA	TGA	ACC	CGTG	
121				-+-			+				+			-+-						GCAC	180
	Е	С	E	R	v	Α	н	I	т	P		G			F		D	E	P	v	_
181				-+-			+				+			-+-			+			GAAT	240
	TT	ccc	CGC	TÄA	GTG	CTA	GAG	GGC	TCI	GTT	GCG	GGT	CTT	TTG	CAA	.CAG	AAA	CGC	TTA	CTTA	
	K	G	R	F	T	I	s	R	D	N	A	Q	K	T	L	S	L	R	M	N	-
	AG	ፐርጉ	GAG	GCC	TGA	GGA	CAC	agI	CGT	'GTA	TTA	CTG	TGC	GGC	AGA	TTG	GAA	ATA	CTG	GACT	
241				-+-			+				+			-+-			+			+	300
	TC	AGA	CIC	CGG	ACT	CCI	GIG	CCG	GCA	CAT	AA1	GAC	ACG	CCG	TCT	AAC	CII	TAI	GAC	CTGA	
	s	L	R	P	E	D	T	A	V	Y	Y	С	A	A	D	W	K	¥	W	T	-
	TG	TGG	TGC	CCA	GλC	TGG	AGG	ATA	CTI	'CGG	ACA	GTG	GGG	TCA	GGG	GGC			CAC	CGTC	
301				-+-			+				+			-+-			+			+ GCAG	360
																				v	_
	С	G	A	Q	T	G	G	Y	F	G	Q	W	G	Q	G	A	Q	V 	_	V	_
									LCGA										TTC		
361							_													416	
	ΑG	GAG	TGA	TCG	ATC	LAA	GGG	CAI	GC1	'GCA	AGG	CCI	GAT	GCC	AAG	AAT	TAT	CTT	AAG	410	

Fig.5C.

,	CA	GGT	GAA	ACT	GCT	OI CGA	GTC	TGC	GGG	AGG	GTC	GGI	GCA	GGC	TGC	AGG	GTC	TCT	GAG	ACTC	60
1	GT	CCA	CTT	TGA	.CGA	GCT	CAG	ACC	ccc	TCC	CAG	CCA	CGI	ccc	ACC	TCC	CAG	AGA	CTC	TGAG	60
	Q	V	K	L	L	E	Ş	G	G	G	S	v	Q	A	G	G	s	L	R	L	-
61				-+-			+				+			+-						GGCT +	120
	AG	GAC	ТТА	ACA	GAG	ACC	GAC	AGC	СТС	ATC	ATG	TAA	'AAC	:GGA	CCC	GAC	CAA	.GGC	:GGT	CCGA	
	S	С	N	ν	s	G	S	P	S	S	T	Y	С	L	G	W	I.	R	Q	A	-
121	cc	AGG	GAA	GGA -+-	GCG	TGA	.GGG	GGT	CAC	AGC	GAT	TAA	CAC	TGA	TGG	CAG	TGT	CAT	ATA	CGCA	180
	GG	TCC	CTT	CCT	CGC	ACT	ccc	CCA	GTG	TCG	CTA	TTA	GTG	ACI	'ACC	GTC	ACA	GTA	TAT	GCGT	
	P	G	ĸ	E	R	E	G	v	T	Α	I	N	T	D	G	S	v	I	Y	A	-
181				-+-			+				+			-+-			+			ATAT +	240
	CG	GCT	GAG	GCA	CTT	CCC	GGC	AAT	GTG	GTA	GAG	GGT	TCT	GTG	GCG	GTT	CTT	TTG	CCA'	TATA	
	A	D	S	v	K	G	R	F	T	I	S	Q	D	T	A	К	IC	T	v	Y	
241	CT	CCA	GAT	GAA -+-		CCT	GCA	ycc	TGA	GGA	TAC	GGC	CAC	CTA	TTA	CTG	CGC	GGC	AAG.	ACTG	300
	GΛ	GGT	CTA	CTT	GIT	GGλ	CGT	TGG	ACT	CCT	ЛTG	CCG	GTG	GAT	'ANT	G VC	GCG	CCG	TTC	TGAC	
	L	Q	M	И	N	L	Q	P	E	D	T	λ	T	Y	Y	С	У	λ	R	L	-
301	AC	GGA	GAT	GGG	GGC	TTG	TGA	TGC	GAG	ATG	GGC	GAC	CTT	AGC	GAC	AAG	GAC	GTT	TGC	GTAT	360
																				CATA	
	T	E	M	G	A	С	D	λ	R	¥	λ	T	L	A	T	R	T	F	A	Y	-
262									GGT		CGI									CGAC	420
361																				GCTG	420
	N	Y	W	G	R	G	T	Q	v	T	v	s	s	L	A	s	Y	P	¥	D	-
422		TCC	GGA	CTA	CGG	TTC	TTA	ATA.	Eco GAA	TTC											
421		AGG	CCT	GAT	GCC	AAG	AAT	TAT	CII		44	9									

HindIII

9 AATCGCCGGCGGTCCACTTTGACGAGCTCATTCACTGATTCCAGGGAGT
A O V K L L E V S S (ECORI) EagI XhoI BATTTAGCGCCCCCAGGTGAAACTGCTCAGAGTGACTAAGGTCACCGTCTCCTCA

HindIII 120 CTTGTTTTTGAGTAGAGTC**TTCTCCTAGACTTAATTA**CTCTTAAGTAGTTTGCCACTATT E Q K L I S E E D L N * * GAACAAAAACTCATCTCAGAAGAGGATCTGAATTAATGAGAATTCATCAAAACGGTGATA EcorI 61

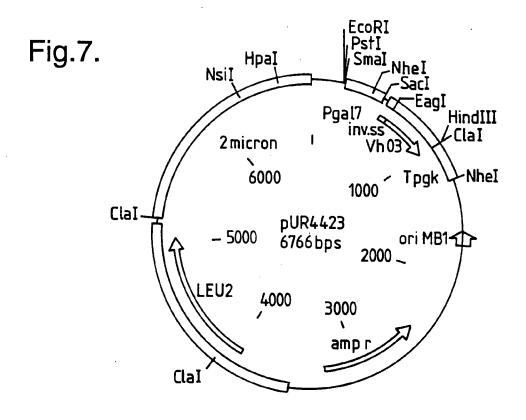
--- 123 121

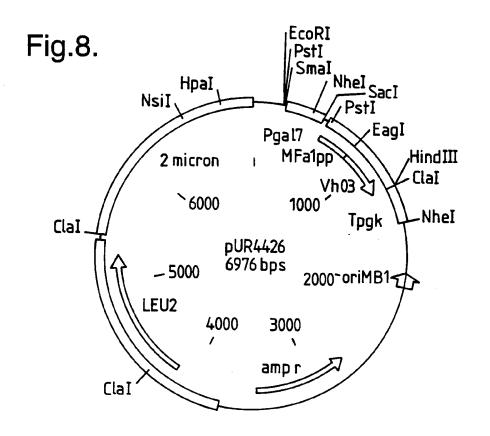
60 AATTTÄGTCGCGACAGGTGAAACTGCTCGAGTAAGTGACTAAGGTCACCGTCTCCTCAGA ATCAGOGCTGTCCACTTGAOGAGCTCATTCACTGATTCCAGTGGCAGAGGAGTCT BSTEII XhoI (ECORI) NruI

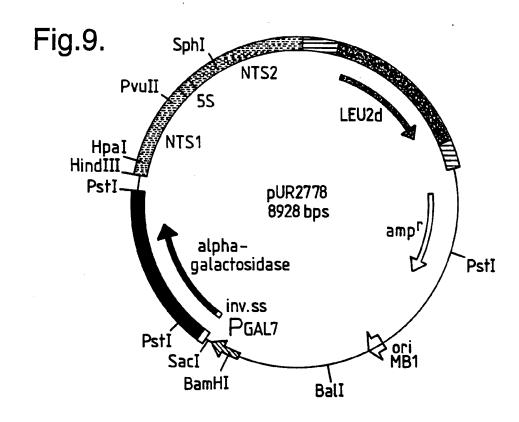
120 tgitttigagtagagtcitctc**ctagacttaa**ttactcttaagtagaattccactattcg Ecori Aflii Hi

61

- 121 A 121







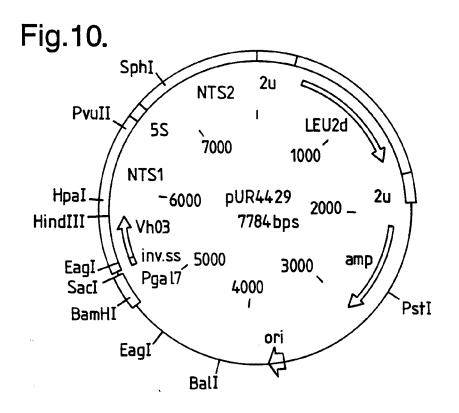
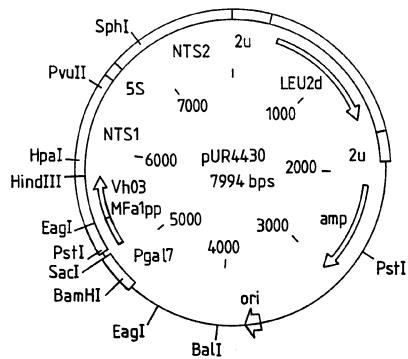
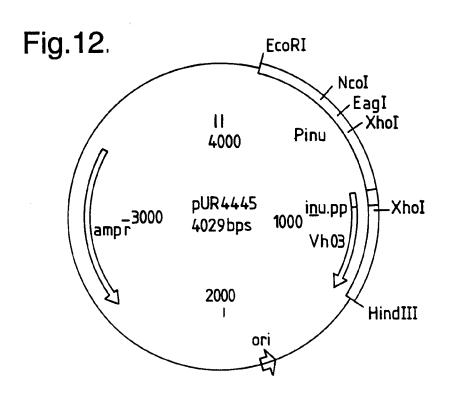
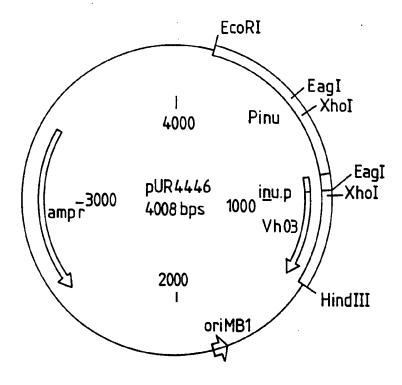


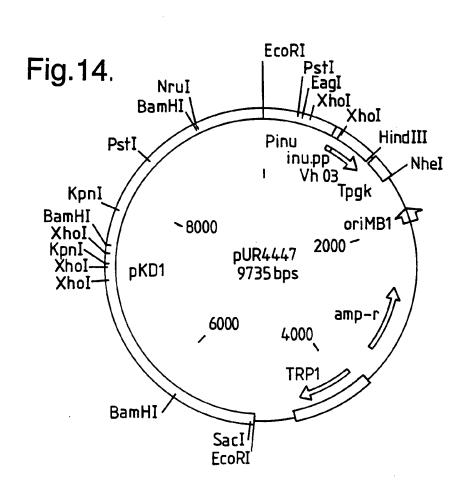
Fig.11.



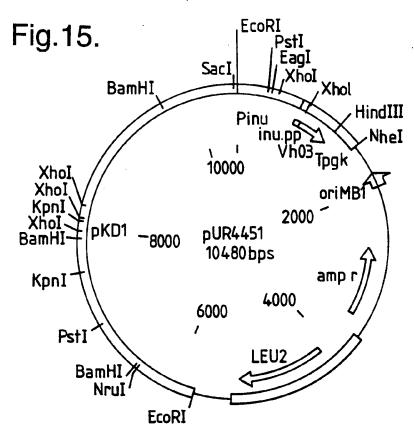


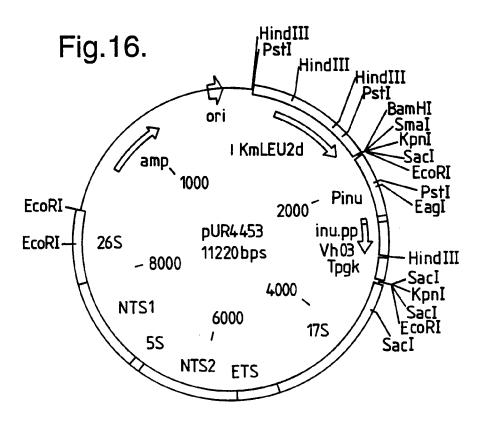












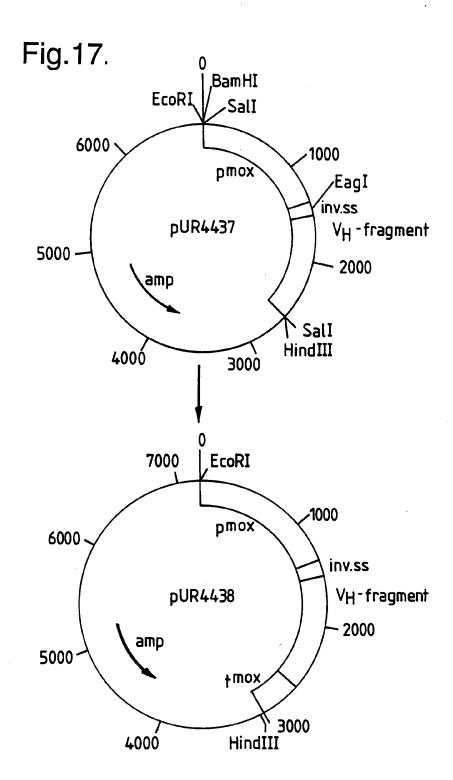
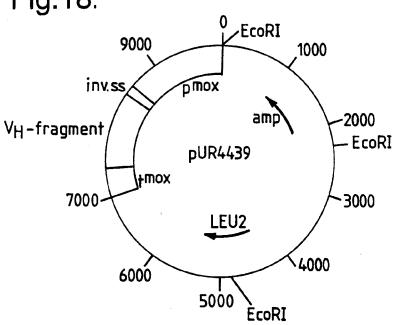


Fig.18.



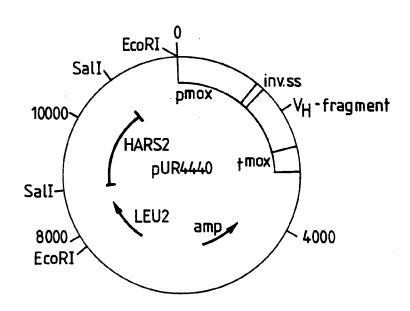
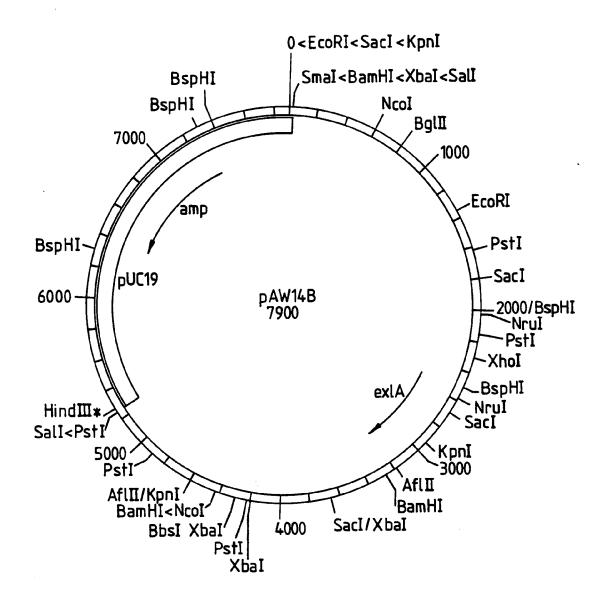


Fig.20.



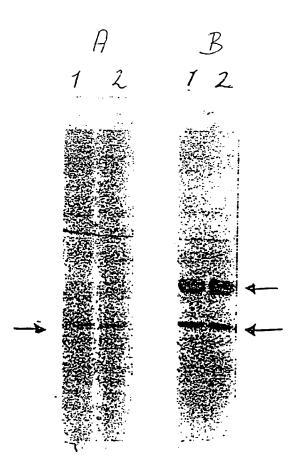


FIGURE 21

INTERNATIONAL SEARCH REPORT

Inter nal Application No PCT/EP 94/01442

A. CLASSIFICATION OF SUBJECT MATTER IPC 5 C12N15/13 C07K15/28 A61K39/395 According to International Patent Classification (IPC) or to both national classification and IPC **B. FIELDS SEARCHED** Minimum documentation searched (classification system followed by classification symbols) C12N C07K A61K IPC 5 Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Electronic data base consulted during the international search (name of data base and, where practical, search terms used) C. DOCUMENTS CONSIDERED TO BE RELEVANT Relevant to claim No. Citation of document, with indication, where appropriate, of the relevant passages Category ' 1,3 EP,A,O 256 421 (PHILLIPS PETROLEUM A COMPANY) 24 February 1988 cited in the application see the whole document 1,4, 10-12 P,X **NATURE** vol. 363, no. 6428 , 3 June 1993 , LONDON, pages 446 - 448 C. HAMERS-CASTERMAN ET AL. 'Naturally occurring antibodies devoid of light chains. cited in the application see the whole document -/--Patent family members are listed in annex. Further documents are listed in the continuation of box C. Special categories of cited documents: "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the "A" document defining the general state of the art which is not considered to be of particular relevance 'E' earlier document but published on or after the international "X" document of particular relevance; the claimed invention filing date cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled document referring to an oral disclosure, use, exhibition or other means document published prior to the international filing date but later than the priority date claimed "&" document member of the same patent family Date of mailing of the international search report Date of the actual completion of the international search 26 -08- 1994 19 August 1994 Authorized officer Name and mailing address of the ISA European Patent Office, P.B. 5818 Patentiaan 2 NL - 2280 HV Rijswijk Tel. (+ 31-70) 340-2040, Tx. 31 651 epo nl, Nooij, F Fax: (+31-70) 340-3016

INTERNATIONAL SEARCH REPORT

Inte mal Application No
PCT/EP 94/01442

	•	PCT/EP 94	1/01442
	ation) DOCUMENTS CONSIDERED TO BE RELEVANT		Relevant to claim No.
Category °	Citation of document, with indication, where appropriate, of the relevant passages	: =	Relevant to claim No.
P,X	FEBS LETTERS vol. 339, no. 3 , 21 February 1994 , AMSTERDAM, THE NETHERLANDS pages 285 - 290 J. DAVIES ET AL. ''Camelising' human antibody fragments: NMR studies on VH domains.' see the whole document		1,5, 10-12
Ρ,Χ	WO,A,94 04678 (C. CASTERMAN ET AL.) 3 March 1994 see the whole document		1,3,4,6,10-12

INTERNATIONAL SEARCH REPORT

...ormation on patent family members

Inter nal Application No
PCT/EP 94/01442

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP-A-0256421	24-02-88	AU-B- 620667 AU-A- 4590789 AU-B- 594476 AU-A- 7474787 JP-A- 63044899	20-02-92 22-03-90 08-03-90 18-02-88 25-02-88
WO-A-9404678	03-03-94	EP-A- 0584421 AU-B- 4949793	02-03-94 15-03-94

Form PCT/ISA/210 (patent family annex) (July 1992)